MOUNT SINAI
GUSTAVE L.
AND
JANET W.
LEVY
LIBRARY

MOUNT SINAI
GUSTAVE L.
AND
JANET W
LEVY
LIBRARY

Digitized by the Internet Archive
in 2015

https://archive.org/details/mountsinaijourna7720moun

MOUNT SINAI JOURNAL OF MEDICINE

Journal of Translational and Personalized Medicine

VOL. 77, NO. 1
JANUARY/FEBRUARY 2010

Issue Theme
Alzheimer's Disease: New Light on Etiology and Management

Theme Editor
Nikolaos Robakis, MD

irrent Treatment and
ecent Clinical Research
Alzheimer's Disease

ccination as a
erapeutic Approach

europathology

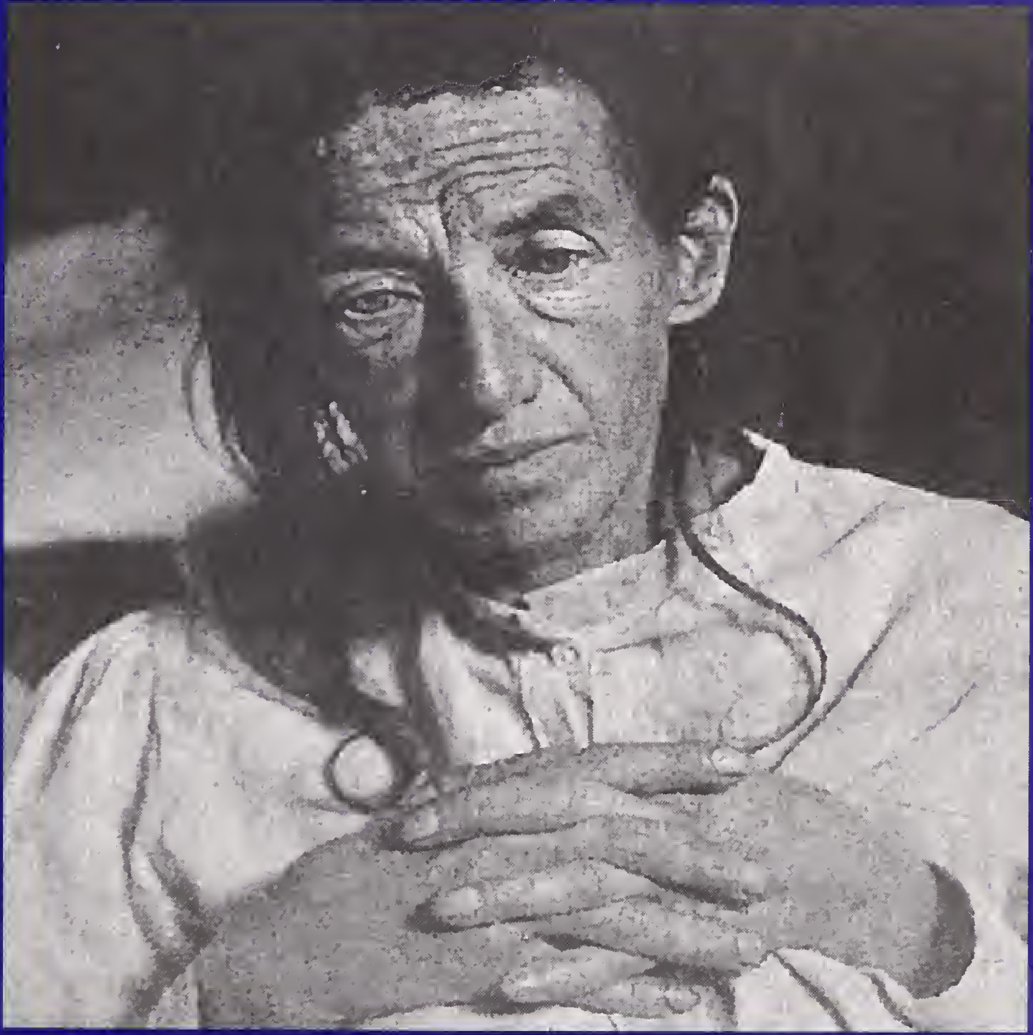
nyloid-ß Oligomers:
ssible Role as
y Neurotoxins

lation between
u Pathology and
euroinflammation

ll "Self-eating"
utophagy) Mechanism

nsgenic Mouse Models

scular Risk Factors and
scular Dysfunction



Auguste D, First Patient Diagnosed with Alzheimer's Disease (1906)

pecial Features

nslating Knowledge about Environmental Health to Pediatricians
rgical Treatment of Metastatic Spinal Tumors
e First Season of 2009 H1N1 Influenza

WILEY-BLACKWELL

Recommend this Journal
to Your Library Today!

MOUNT SINAI
JOURNAL OF
MEDICINE

A Journal of Translational and Personalized Medicine

Editor-in-Chief: **Penny A. Asbell, MD, FACS, MBA**
Mount Sinai School of Medicine, New York, NY

Continuously published since 1934 and included in the *Index Medicus* from the very first volume, the **Mount Sinai Journal of Medicine (MSJM)** has a distinguished career as a general medical journal for a worldwide audience.

Dedicated to improving communication among all participants in medicine today, **MSJM** is an excellent source of up-to-date medical information for healthcare clinicians, educators, researchers and policy makers worldwide. **MSJM** will inspire the creative thinking needed to create new ways to understand, prevent, and treat human disease.

Each issue includes several quality, in-depth review articles related to a single theme and several others on subjects of diverse clinical interest. All theme articles are written by renowned authorities in their fields, from both Mount Sinai and elsewhere, under the supervision of Theme Editors.

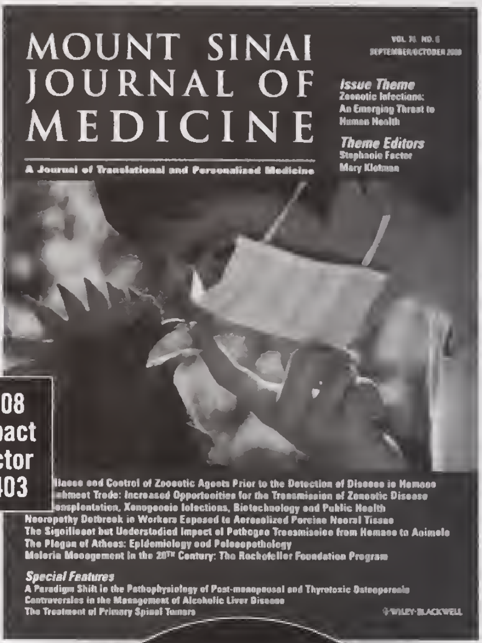
Upcoming Themes for Volume 77:

- Alzheimer's Disease: New Light on Etiology and Management
- Genes, Race, Ethnicity and the Environment
- Endovascular Surgery: Advances in Science and Technique
- Drug Discovery in Academic Settings
- The Obesity Pandemic
- Children's Health and the Environment

ACCESS THESE 2009 EDITORS CHOICE ARTICLES FREE ONLINE

- Advances and Controversies in Assisted Reproductive Technology (76:6)
- Xenotransplantation, Xenogenic Infection, Biotechnology and Public Health (76:5)
- The Student-run Health Clinic: A Novel Arena to Educate Medical Students on Systems-based Practice (76:4)
- Glycemic Targets for Patients with Type 2 Diabetes Mellitus (76:3)
- Neurobehavioral Sequelae of Traumatic Brain Injury (76:2)
- Spectrum of Fibrosing Diffuse Parenchymal Lung Disease (76:1)

Want more? Access a free sample issue online today!



MOUNT SINAI
JOURNAL OF
MEDICINE

A Journal of Translational and Personalized Medicine

Issue Theme
Zoonotic Infections:
An Emerging Threat to
Human Health

Theme Editors
Stephanie Factor
Mary Klotman

Neuropathy Outbreak in Workers Exposed to Aerosolized Porcine Neural Tissue
The Significant but Understudied Impact of Pathogen Transmission from Humans to Animals
The Plague of Athens: Epidemiology and Paleopathology
Malaria Management in the 20TH Century: The Rockefeller Foundation Program

Special Features
A Paradigm Shift in the Pathophysiology of Post-menopausal and Thyrotoxic Osteoporosis
Controversies in the Management of Alcoholic Liver Disease
The Treatment of Primary Spinal Tumors

WILEY-BLACKWELL

2008
Impact
Factor
1.403

Stay on Top
of the Research!

Sign-up for Email Table of Contents Alerts
delivered direct to your desktop – FREE!

Simply go to the homepage and click on
"Set E-Mail Alert" for automatic
table of contents alerts,
or create a customized alert
based on your own
search criteria.

How to Receive Complimentary Online Access for Institutions?

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Library Recommendation Form

To________________________________

Date______________________________

Requested by:_______________________

Dept.______________________________

I have reviewed the **Mount Sinai Journal of Medicine** and urge you to add a complimentary 2010 online subscription to our subscription list. Please contact me should you have any questions.

WWW.MOUNTSINAIJOURNAL.ORG

WILEY-BLACKWELL

Enter Your 2010 Subscription Today!

MOUNT SINAI JOURNAL OF MEDICINE
A Journal of Translational and Personalized Medicine
Issue Theme
Type 2 Diabetes Mellitus: New Evidence, More Controversy
Theme Editor
Dina Green
Special Features
WILEY-BLACKWELL

Aimed at clinicians and educators, the **Mount Sinai Journal of Medicine** provides high-quality, in-depth review articles in every issue, studded with clinical pearls.

Your 2010 will include the six 2010 print issues of the **Mount Sinai Journal of Medicine** (Volume 77), plus concurrent full-text online access to these six issues on Wiley's Internet publishing platform.

Standard Personal Subscriptions

$130.00 (U.S./Canada/Mexico)
$120.00 (Outside North America)

Mount Sinai Alumni Subscriptions!

$40.00 (Worldwide)

Extended for Institutions in 2010

Complimentary Online Access!

Your institution can opt to receive complimentary online access to the **Mount Sinai Journal of Medicine** during 2010. When registering for complimentary online access, your institution can also obtain one complimentary print subscription.

PERSONAL SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 personal subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $105.00 (U.S./Canada/Mexico) ☐ $130.00 (Outside North America)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

Method of Payment (Canada residents please add GST. Personal rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **John Wiley & Sons, Inc.** Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA

Card #____________________ Exp. Date__________

Signature______________________________

(credit card orders not valid without signature)

Ship To:

Name______________________________

Address______________________________

City________________State________Zip__________

MSJM0910

MSJM ALUMNI SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 Alumna/Alumnus Subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $40.00 (Worldwide)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

METHOD OF PAYMENT (Alumna/Alumnus rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **The Mount Sinai School of Medicine**. Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA ☐ Other

Card #____________________ Exp. Date__________

Signature______________________________

(credit card orders not valid without signature)

Ship To:

Name______________________________

Address______________________________

City________________State________Zip__________

To subscribe online visit: **www.msjmonline.com**

MSJM10

COMPLIMENTARY ONLINE ACCESS FOR INSTITUTIONS

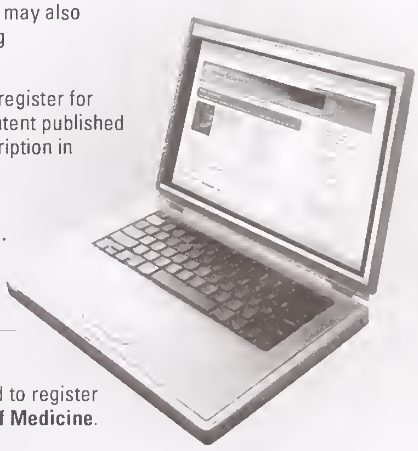
Mount Sinai Journal of Medicine

A Journal of Translational and Personalized Medicine

2010. Volume 77. 6 issues per year. Print ISSN: 0027-2507. Online ISSN: 1931-7581.

Your institution can opt to receive complimentary online access to the **Mount Sinai Journal of Medicine** during 2010. Your institution may also request one complimentary print subscription when registering for complimentary online access.

There is no obligation to subscribe in subsequent years if you register for complimentary access. Institutions retain online access to content published during the complimentary period, with or without a paid subscription in subsequent years.



Dear Librarian,
I have reviewed the scope of Mount Sinai Journal of Medicine.
Please register for complimentary access:

Name: ________________ Department: __________

Signature: ________________ Date: __________

Ask your librarian to e-mail **optinaccess@wiley.co.uk** or visit **www.interscience.wiley.com/msjregister** for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Don't delay, recommend to your library today!

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO. 2277 HOBOKEN, NJ

POSTAGE WILL BE PAID BY ADDRESSEE

JOHN WILEY & SONS INC
ATTN: BETH MOLD, MAIL STOP 8-01
111 RIVER STREET
HOBOKEN, NJ 07030-9441

PLACE
POSTAGE
STAMP
HERE

INES SANTOS
MOUNT SINAI JOURNAL OF MEDICINE
ONE GUSTAV L. LEVY PLACE BOX 1094
NEW YORK, NY 10029-6574

Library Routing Card

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

Editor-in-Chief

Penny A. Asbell

Deputy Editor

David H. Adams

Associate Editors

Stuart Aaronson
Erwin Bottinger
Dennis S. Charney
Valentin Fuster
Wayne Goodman
Jonathan Halperin
Ravi Iyengar
Mary Klotman
Paul Klotman
Ihor R. Lemischka
Michael L. Marin
Thomas G. McGinn
Ira Nash
Maria New
Rosamond Rhodes
Hugh Sampson
Mary Sano
Jeffrey H. Silverstein
Albert Siu
Anthony Squire
Fred Suchy
Savio Woo

Editorial Board

William J. Bremner	University of Washington
Thomas M. Coffman	Duke University
Dwight Evans	University of Pennsylvania
Gabriel Haddad	University of California, San Diego
Heidi E. Hamm	Vanderbilt University
Matti Häyry	University of Manchester
Eve J. Higginbotham	Morehouse School of Medicine
Katherine A. High	University of Pennsylvania
Tony Hope	University of Oxford
John H. Krystal	Yale University
Raju Kucherlapati	Harvard Medical School
Robert J. Lefkowitz	Duke University
Robert Lisak	Wayne State University
George Lister	University of Texas
Philip A. Marsden	University of Toronto
Charles Nemeroff	Emory University
Scott Rauch	Harvard Medical School
Bernard Roizman	University of Chicago
Alan F. Schatzberg	Stanford University
Inder Verma	The Salk Institute
Jeffrey A. Whitsett	University of Cincinnati
Allan Wolkoff	Albert Einstein College of Medicine
Noam Zohar	Bar-Ilan University

Theme Editor

Nikolaos K. Robakis, Ph.D.
First AP Slaner Professor for Alzheimer Disease Research
Director, Center for Molecular Biology and Genetics of Neurodegeneration
Mount Sinai School of Medicine, New York, NY

Editorial Director

Stephen Cohen
stephen.cohen@mssm.edu
212-241-6304

Publisher

John Wiley & Sons
Anna Wesley, Associate Editor
awesley@wiley.com

Auguste D, First Patient Diagnosed with Alzheimer's Disease

At a scientific meeting in Frankfurt, Germany in November 1906, Dr. Alois Alzheimer presented the case of "Frau Auguste D.," a 51-year-old woman brought to Dr. Alzheimer by her family in 1901. Auguste's family told him she had had problems with memory, difficulty speaking and understanding what was said to her, and unfounded suspicions that her husband was unfaithful. Her symptoms rapidly grew worse, and within a few years she was bedridden. She died in the spring of 1906, the first patient to be diagnosed with the illness that became known as Alzheimer's disease.

Future *MSJM* Themes and Theme Editors

- Obesity (Robert Yanagiswa, Derek LeRoith)
- Drug discovery in academic settings (Ravi Iyengar)
- Genes, race, ethnicity and the environment (Gary Butts)
- Parkinson's Disease (Michele Tagliati, Ron Alterman)
- Children's health and the environment (Philip Landrigan)
- Endovascular surgery (Michael Marin)

©2010 **Mount Sinai School of Medicine.** All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, Phone: (978) 750-8400; Fax: (978) 750-4770. Such requests and other permission inquiries should be addressed to the Permissions Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; Phone: (201) 748-6011; Fax: (201) 748-6008; E-mail: permreq@wiley.com. Printed in the United States of America by The Sheridan Press.

Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine (Print ISSN: 0027-2507; Online ISSN: 1931-7581 at Wiley InterScience, www.interscience.wiley.com) is published twelve times per year for the Mount Sinai School of Medicine by Wiley Subscription Services, a Wiley Company, 111 River Street, Hoboken, NJ 07030.

Send subscription inquiries to: Subscription Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-6645.

Offprint sales and inquiries should be addressed to the Reprint Billing Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-8776. All other inquiries should be addressed to the Customer Service Department at (201) 748-6645.

Individual subscription price: $95.00 in North America, $120.00 outside North America. Payment must be made in US dollars drawn on a US bank. Claims for undelivered copies will be accepted only after the following issue has been delivered. For more subscription prices and information, please visit www.mountsinaijournal.org.

Change of address: Please forward six weeks prior to move: present mailing label with change of address. Please enclose a copy of the mailing label or cite your subscriber reference number in order to expedite handling. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits. Send claims in care of: *Mount Sinai Journal of Medicine*, c/o John Wiley & Sons, Inc., Attn: Journals Admin Dept. UK, 111 River Street, Hoboken, NJ 07030.

Cancellations: Subscription cancellations will not be accepted after the first issue has been mailed. Periodicals postage paid at Hoboken, NJ and at additional mailing offices.

Postmaster: Send address changes to *Mount Sinai Journal of Medicine*, Subscription Distribution, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Printed in the United States of America. Single issues from current and recent volumes are available at the current single issue price from cs-journals@wiley.com. Earlier issues may be obtained from Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. Tel: +1 518 537 4700, Fax: +1 518 537 5899, Email: psc@periodicals.com

Advertising inquiries should be addressed to: Advertising Sales Manager, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Phone: (201) 748-6921; Fax: (201) 748-6207. Advertising in this journal is published with a caveat emptor understanding. The Publisher, Editors, and members of the Editorial Board do not imply warranty of products, nor quality, validity, or approval of education material offered by such advertising.

Indexed by: Biological Abstracts® (Thomson ISI), Chemical Abstracts Service/SciFinder (ACS), Current Contents®/Clinical Medicine (Thomson ISI), EMBASE/Excerpta Medica (Elsevier), Index Medicus/MEDLINE/PubMed (NLM), Science Citation Index Expanded™ (Thomson ISI), SCOPUS (Elsevier), Web of Science® (Thomson ISI).

This journal is printed on acid-free paper.

Disclaimer: All articles published, including but not limited to original research, reviews, clinical notes, clinical pearls, prismatic cases, clinician's corner, editorials, reports, and letters, represent the opinions and views of the authors and do not reflect any official policy or medical opinion of the Mount Sinai School of Medicine, the institutions with which the authors are affiliated, or the Publisher unless this is clearly specified. Articles published herein should not be relied upon as recommending or promoting a specific method, diagnosis, or treatment by physicians for any particular patient or disorder. Although the Editors and Publisher believe that drug selections and dosages and the specifications and usage of equipment and devices as set forth herein are in accord with current recommendations and practice at the time of publication, they accept no legal responsibility for any errors or omissions, and make no warranty, express or implied, with respect to material contained herein.

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

CONTENTS
VOL 77, NO 1

Theme Introduction

1 Alzheimer's Disease: New Light on Etiology and Management
Nikolaos Robakis

Theme Articles

3 Current Treatment and Recent Clinical Research in Alzheimer's Disease
Judith Neugroschl, Mary Sano

17 Vaccination as a Therapeutic Approach for Alzheimer's Disease
Thomas Wisniewski, Allal Boutajangout

32 Neuropathology of Alzheimer's disease
Daniel P. Perl

43 Amyloid-ß Oligomers: Possible Roles as Key Neurotoxins in Alzheimer's Disease
Alex L. Lublin, Sam Gandy

50 Relationship between Tau Pathology and Neuroinflammation in Alzheimer's Disease
Maria Jose Metcalfe, Maria E. Figueiredo-Pereira

59 Cell "Self-eating" (Autophagy) Mechanism in Alzheimer's Disease
Sarah F. Funderburk, Bridget K. Marcellino, Zhenyu Yue

69 Transgenic Mouse Models of Alzheimer's Disease
Gregory A. Elder, Miguel A. Gama Sosa, Rita De Gasperi

82 Role of Vascular Risk Factors and Vascular Dysfunction in Alzheimer's Disease
Dara L. Dickstein, Jessica Walsh, Hannah Brautigam, Steven D. Stockton Jr., Samuel Gandy, Patrick R. Hof

Special Features

103 First Season of 2009 H1N1 Influenza
Kristin A. Swedish, Gina Conenello, Stephanie H. Factor

114 Translating Knowledge About Environmental Health to Practitioners: Are We Doing Enough?
Leonardo Trasande, Nicholas Newman, Linda Long, Genevieve Howe, Beth J. Kerwin, Richard J. Martin, Sheila A. Gahagan, William B. Weil

124 Surgical Treatment of Metastatic Spinal Tumors
Scott A. Meyer, Harshpal Singh, Arthur L. Jenkins

SUBMIT YOUR WORK TO

Mount Sinai Journal of Medicine

The *Mount Sinai Journal of Medicine* state-of-the-art online submission system allows you to track your paper through the entire production process. You will never have to wait, or wonder, when your paper will publish.

Submission is fast, easy, and safe.
Follow these 5 simple steps:

LOG ON TO
http://mc.manuscriptcentral.com/msjm
Log on with your user ID and password. Click "Author Center" and then "Submit First Draft of a New Manuscript."

TITLE, ABSTRACT & KEYWORDS
Copy and paste these from your document. Use the "Character Palette" to add Greek and other symbols.

UPLOAD FILES USING THE FILE MANAGER
Use the "Browse" button to select a file. Choose a designation from the list and whether it will be seen by referees. Click "Upload" to add your file. Repeat until all files are uploaded.

VIEW 'PROOF'
You must check the 'proofs' before submission. This will show you what Editors and Reviewers will see. You can go back and replace incorrect files.

SUBMIT AND TRACK YOUR PAPER
Click "Submit" once you are satisfied. An automatic confirmation with your manuscript number will follow and the progress of your submission can be tracked.

SUBMISSION HELP

Online
Select the red "Get Help Now" link in the upper right-hand corner of the login screen.

Phone
434-964-4100 (local)
888-503-1050
(domestic toll-free)

REQUIREMENTS

File Types
All types can be uploaded but not all users will have access to all software. To minimize problems we recommend you upload text as Word or RTF, and figures as JPEG, GIF, TIFF, BMP, PNG, EPS, Powerpoint or Excel.

Where possible combine multiple files into a single document.

Submit at:
http://mc.manuscriptcentral.com/msjm

WILEY-BLACKWELL

THEME INTRODUCTION

Alzheimer's Disease: New Light on Etiology and Management

Nikolaos K. Robakis, PhD

A.P. Slaner Professor for Alzheimer's Disease Research, Professor of Neurobiology,
Center for Molecular Biology and Genetics of Neurodegeneration,
Mount Sinai School of Medicine, New York, NY

Alzheimer's disease (AD) is a progressive neurodegenerative disorder of the central nervous system leading to the most common form of age-associated dementia. More than a century has passed since 1907 when Alois Alzheimer first described the neuropathological hallmarks of the dementia that now bears his name. During this time, our understanding of the biological, environmental, and genetic factors that drive this disorder followed progress in the fields of molecular biology, neuroscience, and genetics. It could be said that the modern era of understanding the biological factors involved in AD started in the early 1980s with the first efforts to elucidate the molecular nature of the neuropathological markers found in the brains of affected individuals. These markers include neuritic amyloid plaques, neurofibrillary tangles, and cerebrovascular amyloidosis. Interestingly, a volume on a symposium on amyloid and amyloidosis held in 1979 makes no reference to the amyloidosis of AD,[1] and this indicates that even at this late date, there was little effort directed at the neuropathology of AD. Neuritic amyloid plaques are complex extracellular structures containing at their core amyloid deposits of fibrillar amyloid beta (Aβ) protein surrounded by reactive astrocytes, microglia, and dystrophic neurites. In contrast to the extracellular localization of neuritic amyloid plaques, neurofibrillary tangles accumulate inside neuronal cells and consist mainly of paired helical filaments of hyperphosphorylated tau protein. Many patients also display high levels of amyloid deposits of Aβ peptides in brain blood vessels; this condition is termed *cerebrovascular amyloidosis*. It is now believed, however, that dementia is caused by extensive neuronal and synapse losses in the hippocampus and neocortical regions of the brain.

AD is a serious health problem: it is estimated that by 2020, 30 million people will be afflicted worldwide. In the United States of America, the average cost of care per patient from diagnosis to death has been estimated to be as high as $174,000. The first clinical symptoms of AD include a loss of recent memory, faulty judgment, personality changes, and a progressive loss of reasoning power. At later stages, patients become agitated, aggressive, and delusional with sleep disturbances and loss of reasoning. Most AD cases occur after the ages of 65 or 70, and death ensues about 10 to 15 years after the appearance of the first symptoms. The vast majority of AD cases are termed *sporadic* because they lack a clear genetic etiology, but about 5% of all cases display clear genetic linkages and are classified as familial Alzheimer's disease (FAD). These usually occur at younger ages and follow a more aggressive clinical course than the sporadic forms of AD. The brain neuropathology, however, is similar in sporadic AD and FAD, and this suggests the involvement of common cellular mechanisms in all AD cases. Despite intense efforts, there is still no chemical test to assay for AD, and because clinical symptoms similar to those of AD may be caused by a number of conditions or diseases, including vascular dementia, a definite diagnosis of AD is possible only after clinical symptoms are combined with a postmortem examination of

Address Correspondence to:

Nikolaos K. Robakis
Center for Molecular Biology
and Genetics of
Neurodegeneration
Mount Sinai School of Medicine
New York, NY
Email: nikos.robakis@exchange.
mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20166

© 2010 Mount Sinai School of Medicine

brain tissue for the detection of plaques and tangles. In this volume, Daniel Perl gives a comprehensive review of the neuropathological and clinical findings that characterize AD.

Although in the last quarter of a century we have learned a great deal about this disease, including the chemical composition of the neuropathological markers of AD and information about the genetics of this disorder, the main mechanisms responsible for the accelerated neuronal cell loss that results in dementia are still poorly understood. It is generally accepted, however, that like the pathogenesis of many other disorders, the pathogenesis of AD is complex, driven by both environmental and genetic factors. Presently, aging and the gene encoding apolipoprotein allele E4[2] are the two largest known risk factors for sporadic AD. In contrast to sporadic AD, FAD is mostly driven by specific genetic mutations localized in at least 3 distinct genes, including those encoding the amyloid precursor protein (APP), Presenilin 1 and Presenilin 2. APP is important to all forms of AD because in addition to its specific involvement in the development of FAD, APP is the precursor of the Aβ peptides that aggregate to form the deposits of amyloid fibrils used to define this disorder. APP is also important to the neuropathology of Down syndrome because patients over the age of 40 years develop amyloid deposits identical to those found in patients with AD. Localization of the APP gene on chromosome 21 revealed a direct genetic linkage between these two disorders.[3] A common theory posits that amyloid deposits of Aβ fibrils or soluble oligomeric forms of Aβ peptides are the main causes of the neurodegeneration of AD.[4] There are several weaknesses to this theory however, including studies that show no significant correlation between brain amyloid loads and degree of dementia and data indicating that soluble Aβ species become toxic only at concentrations ten of thousand times higher than their *in vivo* concentrations (for recent review of involvement of Aβ and derivatives in AD see ref. 5). There are several weaknesses to this theory, however, including studies that show no significant correlation between the brain amyloid loads and the degree of dementia. Furthermore, soluble Aβ peptides are normal components of human serum and cerebrospinal fluid, and presently there is little evidence of disease-associated changes in soluble Aβ or its oligomeric forms. In this volume, Gandy and Lublin give a critical review of the involvement of soluble Oligomers in the development of AD.

Recent findings indicate that dysfunctions of the cerebrovascular system caused by cerebrovascular amyloidosis, oxidative stress, and genetics may play important roles in the pathogenesis of AD. Dickstein and colleagues offer an extensive analysis of factors that may promote cerebrovascular abnormalities and their potential involvement in AD. A recent theory suggests that the cellular process of autophagy, which functions to eliminate cellular waste, may also be involved in AD. In this volume, Funderburk *et al.* explore this topic and offer an insightful description of the relationship between autophagy and neurodegeneration. Neuroinflammation and neurotoxicity have also been associated with AD, and many groups are actively examining the potential contribution of these processes to neuronal cell death in AD. Metcalfe and Pereira review this important topic in this volume. Transgenic animal models greatly facilitate studies of human disorders, and this technology has been used extensively to probe factors involved in neuronal dysfunction and behavioral abnormalities. Elder and colleagues cover this topic here.

The most important issue for AD patients and their families is therapeutic intervention. There is a clear need for more efficient therapeutics as currently available treatments provide largely symptomatic relief with only minor effects on the course of the disease. To this end, modern science employs an array of approaches, including immunomodulators, inhibitors of Aβ peptides, vaccination, anti-tau aggregation agents, neuroprotective factors, and anticholinergic agents. Two groups, one headed by Tom Wisniewski and the other by Mary Sano, review current efforts for the development of new therapeutics for AD.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Glenner G. A retrospective and prospective overview of the investigations on amyloid and amyloidosis – the β fibrilloses. In: Gleener GG, Costa PP, Freitas FD, eds. *Amyloid and Amyloidosis*. Amsterdam, the Netherlands: Excerpta Medica; 1980; 3–13.
2. Corder EH, Saunders AM, Strittmatter WJ, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 1993; 261: 921–923.
3. Robakis NK, Wisniewski HM, Jenkins EC, et al. Chromosome 21q21 sublocalization of the gene encoding the β-amyloid peptide present in cerebral vessels and neuritic (senile) plaques of people with Alzheimer disease and Down syndrome. *Lancet* 1987; Lancet i: 384–385.
4. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* 2002; 297: 353–356.
5. Robakis NK. Are Aβ and its Derivatives Causative Agents or Innocent Bystanders in AD? *Neurodegenerative Dis* 2009 [Epub ahead of print].

DOI:10.1002/MSJ

Current Treatment and Recent Clinical Research in Alzheimer's Disease

Judith Neugroschl, MD,[1] and Mary Sano, PhD[1,2]

[1]Mount Sinai School of Medicine, New York, NY
[2]James J. Peters Veterans Affairs Medical Center, Bronx, NY

OUTLINE

Current Food and Drug Administration-approved Medications
- **Cholinesterase Inhibitors**
- **Memantine**
- **Anti-Inflammatory Agents**
- **Antioxidants**
- **Cholesterol Lowering**
- **Estrogen**
- **Docosahexanoic Acid**
- **Ginkgo Biloba**
- **Homocysteine/B Vitamins**
- **Xaliproden (SR57746A)**

Agents currently under investigation
- **Interventions Targeting Amyloid**
- **Advanced Glycation End Products**
- **Latrepirdine (Dimebon)**
- **Nerve Growth Factor**
- **Resveratrol**
- **Targeting Tau**
- **Serotonin Receptors**

Conclusion

ABSTRACT

The transition from either epidemiological observation or the bench to rigorously tested clinical trials in patients with Alzheimer's disease is crucial in understanding which treatments are beneficial to patients. The amyloid hypothesis has undergone scrutiny recently, as many trials aimed at reducing amyloid and plaque have been completed or are in the testing phase. Examples include modulation of the secretases involved in beta amyloid formation, anti-aggregation agents, and immunotherapeutic trials. Other therapies targeting hyperphosphorylated tau and novel targets such as enhancement of mitochondrial function, serotonin receptors, receptor for advanced glycation end products, and nerve growth factor, as well as other strategies, are discussed. A brief review of the current Food and Drug Administration–approved treatments is included. ***Mt Sinai J Med 77:3–16, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: Alzheimer's disease, amyloid, antioxidant, Dimebon, immunotherapy, receptor for advanced glycation end products, research, secretase, tau, treatment.

Altering the course of Alzheimer's disease (AD) could lead to significant pubic health benefits. For example, an intervention that could delay the onset of AD by 2 years would decrease the incidence in such a way that in 50 years there would be nearly 2 million fewer cases than are currently projected.[1] One can imagine that significantly altering the course of the disease would similarly sharply decrease the need for nursing home placement and could help patients remain functional for much longer. The costs associated with taking care of a patient with AD increase with disease severity,[2] so that if patients could be maintained at an earlier stage, it would be beneficial for the patient, the family, and society. In general, the number of patients with AD is thought to be increasing from approximately 5.3 million in 2009 to 11 to 16 million by 2050.[3]

There are only 5 medications approved by the Food and Drug Administration to treat AD. Four of them are acetyl cholinesterase inhibitors (donepezil, galantamine, rivastigmine, and tacrine), and the fifth is the *N*-methyl-d-aspartate antagonist memantine. These medications ameliorate the symptoms and can

Address Correspondence to:

Judith Neugroschl
Mount Sinai School of Medicine
New York, NY
Email: judith.neugroschl@mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20165

© 2010 Mount Sinai School of Medicine

improve the functioning of patients with AD, but they are not curative, nor do they significantly change the course of the illness. The most widely studied treatments aim to address the neuropathological findings over the last century and focus on acetylcholine, inflammatory markers, amyloid plaques, and tau-based neurofibrillary tangles. In this article, we discuss some of the drugs that target these and other novel mechanisms and the rationale for some of the most promising new agents, and we review recent findings, both positive and negative, in the treatment and prevention of AD.

> *There are only 5 medications approved by the Food and Drug Administration to treat AD. These medications ameliorate the symptoms and can improve the functioning of patients with AD, but they are not curative, nor do they significantly change the course of the illness.*

CURRENT FOOD AND DRUG ADMINISTRATION–APPROVED MEDICATIONS

Cholinesterase Inhibitors

Since before 1915, scopolamine, an anticholinergic agent, was used to induce amnesia in women during labor and delivery.[4,5] In the 1970s, research explored the specific cognitive effects of acetylchoine,[6] and trials of physostigmine were undertaken in cognitively normal individuals (eg, Davis *et al.*[7]). Later research was done with AD patients using intravenous and oral physostigmine (eg, Davis and Mohs[8] and Mohs *et al.*[9]). In 1983, the first acetyl cholinesterase inhibitor was approved [tacrine (Cognex)]. It was hampered by 4 times per day dosing, reversible hepatocellular injury in up to 50% of patients, and very poor tolerability due to significant procholinergic side effects of nausea, vomiting, and diarrhea. The other acetyl cholinesterase inhibitors, released between 1996 and 2001, have very similar efficacy, although they have slightly different mechanisms of action. For example, rivastigmine also inhibits butylcholinesterase, and galantamine is an allosteric modulator of the nicotinic receptor. These medications tend to effect a 6- to 18-month improvement in cognitive functioning followed by deterioration along a parallel path. Abrupt cessation of these medications tends to place patients back on their original curve.

Memantine

Memantine is an *N*-methyl-d-aspartate antagonist. It has been approved only for moderate to severe AD, and the Food and Drug Administration refused to extend its approval to mild AD in 2005. In theory, it may help prevent neuronal excitotoxicity. The clinical trials of moderate to severe AD showed a modest improvement in global measures of functioning, both with[10] and without[11] a cholinesterase inhibitor.

Anti-Inflammatory Agents

Overactivity of the immune system, causing damage to neuronal cells, has been postulated to be a mechanism of AD pathogenesis. Neuropathological studies have demonstrated that the brains of AD patients have increased concentrations of acute phase reactants, cytokines, and complement protein in comparison with age-matched controls.[12–15] Epidemiological studies exploring the risk of AD have suggested that nonsteroidal anti-inflammatory or corticosteroid use lessens the risk of developing AD.[16,17] A number of clinical trials have shown no benefit from a variety of anti-inflammatory agents ranging from selective nonsteroidal anti-inflammatory drugs to prednisone as treatments for AD,[18–21] and one recent study has suggested a possible deleterious cognitive effect.[22] Primary and secondary prevention trials with nonsteroidal anti-inflammatory drugs have been similarly negative.[23–25] In addition, there are many side effects of nonsteroidal anti-inflammatory drug use, such as renal insufficiency and gastrointestinal bleeding, which should be noted.

Antioxidants

Oxidative stress has been implicated in cell damage and aging in general[26] and in neurodegenerative diseases in particular.[27] Thus, medications with antioxidant properties have been evaluated as treatments for AD. Many agents have been examined for their potential benefit in dementia and cognitive improvement on the basis of their antioxidant properties. The most common of these is vitamin E. In one study, vitamin E was found to delay clinical progression in AD patients from moderate impairment to severe impairment, although no cognitive benefit was found.[28] Observational studies have reported that the use of high doses of vitamin E may slow progression of AD, and other clinical trials have been unable to demonstrate a benefit in mild cognitive impairment (MCI) or in primary prevention populations.[29,30] There remains considerable interest in vitamin E, and 2 important ongoing studies will be informative. One, sponsored by the Department of Veteran Affairs, is a 3-arm study examining vitamin E (100 IU twice daily), memantine, or placebo in a double-blind fashion in a 4-year study of over 800 subjects with mild AD. The primary outcome is a measure of activities of daily living. Another study, Prevention of Alzheimer's Disease by Vitamin E and Selenium (PREADVISE; ClinicalTrials.gov: NCT00040378), is an add-on to a trial to prevent prostate cancer with 35,533 participants, the Selenium and Vitamin E Cancer Prevention Trial (SELECT), which was halted in 2008 because in an interim analysis there were no significant differences between the groups in prostate cancer prevention, although there was an insignificantly increased risk of prostate cancer in the vitamin E group and an insignificantly increased risk of type 2 diabetes in the selenium-only group.[31] The PREADVISE study continues to follow over 5000 individuals who had been taking vitamin E (400 IU) alone or in combination with 200 μg of selenium or placebo and will determine how exposure to these affected the risk of developing dementia. Although some safety concerns have been raised about the use of high-dose vitamin E in populations at risk for cardiovascular disease, there is no evidence of negative effects on cognition.

Cholesterol Lowering

Initially, epidemiological studies suggested a decreased risk of AD in patients taking 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase inhibitors (statins).[32,33] Basic research supported the idea that statins may inhibit beta amyloid (Aβ) formation in vitro.[34] The initial findings led to clinical research. In a single-site, double-blinded, placebo-controlled, randomized clinical trial of 67 subjects with mild to moderate AD, atorvastatin was shown to have a beneficial effect on cognitive measures in comparison with placebo.[35] When the researchers performed secondary analyses, they found that the benefit was more prominent for individuals with baseline elevated cholesterol levels and milder memory impairments.[36] Simvastatin, which has reasonable central nervous system penetration, was also associated with a decrease in cerebrospinal fluid α- and β-cleaved soluble amyloid precursor protein (sAPP) but not Aβ (42) after 12 weeks of 20 mg daily.[37] When these preliminary data were subjected to a larger multicenter trial lasting for 18 months with 406 patients with AD, statins failed to show a benefit over placebo for simvastain.[38] A more recent meta-analysis of epidemiological studies failed to support the beneficial effect of statins on AD risk.[39] Large multicenter trials of statins for the prevention of cardiovascular disease, looking at cognitive outcomes, have also shown a lack of a beneficial effect on cognition or incident dementia.[40,41]

Estrogen

Beginning in the 1930s, studies have examined the possible role of estrogen in cognition.[42] In the 1990s, there was a great deal of interest in a number of population and case control studies looking at estrogen usage and the risk of AD. Many of these studies (eg, Paganini-Hill and Henderson,[43] Tang *et al.*,[44] and Kawas *et al.*[45]) suggested a decreased risk of AD with estrogen use, although others (eg, Barrett-Connor and Kritz-Silverstein,[46] Barrett-Connor and Edelstein,[47] and Brenner *et al.*[48]) did not. Basic science supported the idea that estrogen might be neuroprotective. Estrogen enhances neuronal growth and connectivity and has been shown to promote neuronal survival (eg, DeVoogd and Nottebohm[49] and Matsumoto and Arai[50]), and it up-regulates nerve growth factor (NGF) messenger RNA and neuronal sensitivity to NGF.[51] Early small open-label treatment studies suggested a positive effect of estrogen on cognition in AD. However, results from randomized clinical trials of conjugated equine estrogen (CEE) in women already diagnosed with mild to moderate AD[52–54] have failed to demonstrate improvement of cognition or a slowed rate of cognitive decline. The duration of follow-up in these studies ranged from 4 to 12 months. There was some evidence of deleterious cognitive effects. The use of CEE to treat AD in postmenopausal women is therefore not recommended. The question of whether estrogen would prevent cognitive loss and AD has also been addressed in prospective trials. Surprisingly, the use of estrogen was associated with both increased cardiovascular risks and increased risk of dementia. For example, in women over 65 years, the use of estrogen or estrogen/progesterone actually increased the odds of converting to MCI or dementia.[55,56] A recent randomized trial of CEE in 180 women, less than 1 year after menopause, who were 45 to 55 years old was designed to answer the question of whether the timing of estrogen exposure (ie, perimenopausal exposure) was important in reducing dementia risk. However, this study demonstrated a trend toward worsening cognition in the treated group.[57] These results make it difficult to imagine any cognitive benefit from exogenous estrogen. Despite these results, the question of whether estrogen in another form may be useful is being considered in studies of raloxifene, a selective estrogen receptor modulator: in a secondary outcome analysis of an osteoporosis trial of 5386 postmenopausal women taking raloxifene or placebo, it was suggested that at 120 mg there was a lower relative risk of any cognitive impairment (relative risk = 0.73, 95% confidence interval = 0.53–1.01).[58] There is a current phase 2 trial (NCT00368459) underway in patients with AD, and it aims to enroll 72 patients at 2 sites.[59] Raloxifene has been associated with an increased risk of venous thromboembolism.

Docosahexanoic Acid

Docosahexanoic acid (DHA) is an omega 3 polyunsaturated fatty acid found in fish and some marine algae. It is a component of synaptic plasma membranes and in animal studies has been shown to perform a number of roles in the brain, including affecting the rate of signal transduction, being neuroprotective, and regulating gene expression. Research has suggested that AD patients have lower plasma DHA levels. A clinical trial in Sweden that was published in 2006 found that in patients with very mild AD, omega 3 fatty acids slowed cognitive decline, as evaluated by the Mini-Mental State Examination (MMSE), over a 12-month period, although no effect was seen in patients with moderate stage disease.[60] There have been 2 recent trials, one in AD and one in elderly adults with a memory complaint. The AD trial was an 18-month phase 3 trial looking at DHA and the progression of AD in 402 patients who had low dietary intake of DHA (ClinicalTrials.gov: NCT00440050). The 2 outcome measures, a cognitive outcome [Alzheimer Disease Assessment Scale–Cognitive (ADAS-Cog)] and a global one (Clinical Dementia Scale–Sum of the Boxes), as well as other behavioral outcome measures did not differ between the placebo and active groups, despite increased levels of DHA in the serum and cerebrospinal fluid. There was no difference between the mild and moderate AD patients, but those without the apolipoprotein E4 allele did have a statistically slower rate of decline with DHA.[61] The other trial was a 6-month trial of DHA versus placebo in 485 elderly patients with a memory complaint. The primary outcome measure was a visuospatial episodic memory test. The active group made significantly fewer errors in comparison with their baseline scores.[61] Another trial looking at high-dose and low-dose DHA and eicosapentaenoic acid in cognitively healthy older adults showed no cognitive benefit over the 26 weeks of the trial.[62]

Ginkgo Biloba

Ginkgo biloba, a Chinese herbal medicine used for its "healthful properties," including effects on memory, was evaluated to see if it decreased the incidence of all-cause dementia in elderly patients.[63] Three thousand individuals, approximately 16% of whom met criteria for MCI at the start of the trial, were followed for a median of 6 years on placebo or ginkgo. Over the duration of the trial, they were evaluated for incident dementia every 6 months. The rates of dementia in general, of AD in particular, and of conversion to dementia from MCI were not statistically different in the 2 groups. They did not report data on specific cognitive measures for individuals in the 2 groups. There was no improvement on any secondary outcomes such as overall morbidity. As pointed out in an editorial concerning this trial, potentially serious side effects of gingko, such as increased bleeding risk, need to be considered in light of the fact that there is no convincing evidence that it has any beneficial effects on cognition or dementia incidence in humans in the forms in which it has been used.[64] A 1-year randomized, placebo-controlled trial of 309 AD patients looking at ginkgo biloba as a treatment for AD showed modest improvement on some measures,[65] although 2 later 6-month trials of 214[66] and 513[67] patients showed no benefit. Overall, the research on ginkgo is at best "inconsistent and unconvincing."[68]

Homocysteine/B Vitamins

The Framingham study and others have noted that elevations in plasma homocysteine are a risk factor for cardiovascular disease[69,70] and cerebrovascular disease[71] as well as AD.[72] For example, a study of 1092 subjects observed that elevated homocysteine was an independent risk factor for AD, and levels over 14 μmol/L almost doubled an individual's risk of developing AD.[72] Theories about how homocysteine might act as a risk factor for AD range from potentiating Aβ production or toxicity to vascular-related effects. To see if lowering homocysteine in patients with AD would change the course of the illness, an 18-month prospective trial in 400 mild to moderate AD patients evaluated the effect of lowering homocysteine. The intervention (folate and vitamins B6 and B12) was quite successful at lowering homocysteine, but it did not slow cognitive decline, and there was no difference between the active and placebo groups in the rate of change of a cognitive measure (ADAS-Cog).[73] This was confirmed in another study that also looked at plasma Aβ levels in 300 patients with strokes; again, despite lowering homocysteine, it did not affect the Aβ levels or cognition.[74]

Xaliproden (SR57746A)

Xaliproden is a compound with has neurotrophic effects in virto. It also antagonizes the 5HT1A receptor[75] which may facilitate glutamatergic and cholinergic neurotransmission.[76] Unfortunately, the results of 2 large 18-month clinical trials of xaliproden (over 1000 patients in each trial) as both monotherapy and adjunctive therapy for mild to moderate AD yielded disappointing results, and the company is not developing the drug further for these indications.[77]

AGENTS CURRENTLY UNDER INVESTIGATION

In this section, we discuss avenues currently under investigation, and we group them either by the mechanism if the mechanism is clear or by the agent itself.

Interventions Targeting Amyloid

Amyloid: A Short Review.

The only known genetic forms of AD are caused by mutations in amyloid precursor protein (APP) or the enzymes that are involved in Aβ formation. APP is coded on chromosome 21. Individuals with Down syndrome (trisomy 21) are noted to universally have the neuropathological hallmarks of AD as they age. APP is cleaved either by α-secretase into a soluble fragment (sAPP) or by β secretase [also called β-site amyloid precursor protein–cleaving enzyme (BACE)] and then by γ-secretase, and this results in fragments that are 37 to 42 amino acids in length. The 42–amino acid fragment is most associated with AD pathology. The only known autosomal dominant forms of AD are due to mutations in either APP or regions that code for proteins involved in the secretases, and they are all associated with elevations in Aβ. Because of this genetic evidence, many scientists postulate that the accumulation of Aβ is the inciting factor in the development of AD. The Aβ fragments can either oligomerize to form soluble toxins or else fibrillize to form plaques. Aβ oligomers may play an important role in AD pathogenesis, as they have been found to be highly toxic to neurons by initiating a cascade of events that contribute to activation of microglia, synaptic degeneration, oxidative injury, and apoptosis. Interventions aimed at reducing amyloid plaque burden by either altering amyloid metabolism through enzyme mediation or maximizing clearance, particularly through immunotherapy, are under study (reviewed by Barten and Albright[78]).

The only known genetic forms of AD are caused by mutations in amyloid precursor protein (APP) or the enzymes that are involved in amyloid beta (Aβ) formation.

Modulating the Formation of Aβ.

If the overproduction or accumulation of Aβ is in fact the crucial event in the cascade that leads to AD, then decreasing its production by the inhibition of BACE or γ-secretase or the enhancement of α-secretase should decrease the pathology and change the course of the illness. Research has been done to identify and modulate the secretases. The greatest concern about inhibiting secretases, which have been highlighted in animal models and early trials, is the other potentially important proteins that they may be involved in processing. Thus, toxicities must be carefully monitored with these agents.

If the overproduction or accumulation of Aβ is in fact the crucial event in the cascade that leads to AD, then decreasing its production by the inhibition of β-site amyloid precursor protein-cleaving enzyme or γ-secretase or the enhancement of α-secretase should decrease the pathology and change the course of the illness.

There are a number of BACE inhibitors that have been described in animal models but have not yet arrived in human trials. A number of approaches, including the creation of novel antibodies to target the BACE cleavage site of APP, have been taken.

The research into the γ-secretases has progressed farther, with 1 medication in phase 3 clinical trials. One concern about the γ-secretases is that they are involved in the cleavage of the Notch transmembrane receptors. In animal models, chronic inhibition is associated with effects in the gastrointestinal system, thymus, and spleen.

The agent in phase 3 trials is LY450139 from Eli Lilly. The phase 2 study[79] of 70 AD patients, lasting 6 weeks, demonstrated relative tolerability and decreased plasma levels of Aβ. Two phase 3 trials that are currently recruiting are looking at different doses in comparison with placebo (ClinicalTrials.gov: NCT00762411 and NCT0059456). Other γ-secretase inhibitors in clinical trials include one from Bristol-Myers Squib (BMS-708 163), which is relatively Notch-sparing according to preclinical testing. There are 2 ongoing clinical trials with this agent (ClinicalTrials.gov: NCT00810147 for patients with mild to moderate AD and NCT00890890 for those with prodromal AD). In addition, Wyeth is testing GSI136 in a phase 1 trial (ClinicalTrials.gov: NCT00718731).

Other agents serendipitously have been found to be secretase inhibitors. For example, tarenflurbil, an anti-inflammatory agent, was found to modulate γ-secretase activity. Basic research suggested cognitive benefit in animal studies. The year-long phase 2 trial of 210 patients with mild to moderate AD demonstrated tolerability, but it did not show a beneficial effect on function or cognition. In a post hoc analysis, some benefit was seen at the highest dose in the patients with the mildest disease[80] An 18-month phase 3 trial with 1649 patients showed no benefit from treatment, and the company (Myriad) discontinued the investigation of the agent.[81] It has been postulated that the oral administration of the medication may not have allowed for adequate γ-secretase inhibition to affect disease.[82]

Blocking the Aggregation of Beta Amyloid into Oligomers or Amyloid Plaque.

The amyloid hypothesis posits that aggregated Aβ is particularly toxic to cells, and thus blocking aggregation and ultimately preventing plaque from forming could change the course of the illness. The most interesting agent in trials is scyllo-inositol, which is currently in phase 2 clinical studies sponsored by Elan Pharmaceuticals (ClinicalTrials.gov: NCT00568776). This agent seems to be directed at the Aβ oligomers: it binds them and prevents synaptic damage. Scyllo-inositol is a small molecule that readily crosses the blood-brain barrier by active transport. A recent negative trial of anti-aggregation agents was 3APS [tramiprosate or homotaurine (Alzhemed from Neurochem)]. The phase 3 trial in the United States was inconclusive; it failed to demonstrate efficacy on the cognitive endpoints in the 18-month trial of 1052 patients. The company then halted the European phase 3 trial before the data were released, and the results of the 2 trials have not been published except in a non–peer-reviewed article on the columetric magnetic resonance imaging results, which suggested less hippocampal shrinkage.[83] Citing the magnetic resonance imaging findings, the company has decided to promote this medication as an over-the-counter medication, Vivimind, which is being put forward as protecting against memory loss.[84]

The amyloid hypothesis posits that aggregated Aβ is particularly toxic to cells, and thus blocking aggregation and ultimately preventing plaque from forming could change the course of the illness.

The phase 3 trial in the United States was inconclusive; it failed to demonstrate efficacy on the cognitive endpoints in the 18-month trial of 1052 patients.

Increasing Clearance of Beta Amyloid with Immunotherapy.

Specific antibodies can bind Aβ and promote its clearance from the central nervous system. This can be accomplished either by the introduction of Aβ as an antigen so that the body can produce antibodies to it (active immunotherapy) or by the use of preformed antibodies (passive immunotherapy.) In 2002, the first phase 2 active immunity trial (AN-1792) was halted because of meningoencephalitis in 6% of the patients who received the vaccine (ClinicalTrials.gov: NCT00021723). This trial of 372 patients used an aggregated amyloid peptide with an adjuvant (QS-21) to boost the immune response. Researchers postulated that T cell–mediated immunity caused this adverse reaction, and thus the second-generation agents currently in trials use an N-terminal fragment of Aβ as an antigen because it is thought that this will generate more B cell–mediated, humoral immunity. The T cell–stimulating epitopes are more distal. Despite the halting of the trial, the researchers did follow the patients who received the vaccine. It is interesting to note that less than 20% of those who received the active vaccine achieved the targeted immune response level.[85] In those that did, there was a marked effect on amyloid in the brain, although the other neuropathological hallmarks of AD were unchanged in comparison with controls, with prominent tangles, neuropil threads, and cerebral amyloid angiopathy.[86] Unfortunately, there was little evidence of significant cognitive benefit as measured by both the ADAS-Cog and MMSE. However, with a composite statistical measure on the Neuropsychological Test Battery, there was some suggestion of benefit. With respect to functional measures, the patients who mounted the antibody response and the placebo controls declined at similar rates. There was some suggestion of improvement on a quality-of-life measure in the active group.[85]

Given the significant changes in plaque burden seen on autopsy, there was speculation about why there was little cognitive and functional change seen in the patients who mounted an antibody response. First, it is possible that removing Aβ after AD has already progressed is not useful; rather, it must not be allowed to accumulate in the first place. A second possibility is that Aβ is not the central player that it is thought to be, and other targets might be more fruitful. Third, the small number of patients who responded to the vaccine may limit the trial's generalizability. Had the patients responded to the vaccine safely and had the full set of booster injections been used, it is possible that more significant differences would have emerged in the active and placebo groups. Second-generation antigens are currently in phase 2 trials [eg, CAD106 (Novartis; ClinicalTrials.gov: NCT00795418) and PF-04 360 365 (Pfizer; ClinicalTrials.gov: NCT00945672)].

The approach of using preformed antibodies is also being studied (passive immunization). Passive immunity does not depend on the patient mounting an adequate immune response, and antibodies are cleared within hours of an infusion.[87] Pooled human antibodies (eg, intravenous immunoglobulin, which is purified human immunoglobulin G) have been shown to contain antibodies to Aβ and possibly to Aβ oligomers as well. Recently, a small open-label study was published that followed 8 patients throughout 18 months of treatment.[88] The primary outcome measure was the MMSE. Six of the 8 patients had their scores stabilize or improve at the 18-month mark. Currently, a phase 3 trial (ClinicalTrials.gov: NCT00818662) is underway with a targeted enrollment of 360 patients, and it is trying to determine if intravenous immunoglobulin significantly slows the rate of decline in patients with mild to moderate AD. The primary outcomes are the ADAS-Cog and the Clinical Global Impression of Change scores at 9 months, with a planned secondary analysis at 18 months.

It is also possible to use preformed monoclonal antibodies to Aβ. There are currently 4 phase 3 trials of bapineuzumab, 2 with apolipoprotein E4 carriers and 2 without them. Their target enrollment is over 1000 patients. The results of the 18-month phase 2 trial of 240 patients with AD did not show efficacy on the cognitive measures, but a post hoc analysis looking at patients who did not carry the apolipoprotein E4 allele showed statistically significant benefits on ADAS-Cog and the Neuropsychological Test Battery.[87,89] Of note, the higher doses of the antibodies were associated with vasogenic edema, and thus the current phase 3 trial does not include the highest dosage of bapineuzumab used in phase 2.[90] It is important to note that cerebral microhemorrhages have also been described even with passive immunotherapy.[91]

Advanced Glycation End Products

Advanced glycation end products (AGEs) are formed endogenously during glycation and can also be ingested in a variety of foods.[92] These AGEs have been implicated in aging through a variety of mechanisms, including increased protein crosslinking and increased free radical formation, and as proinflammatory mediators. Receptor for advanced glycation end products (RAGE) is an immunoglobulin supergene family expressed on the cell surface of multiple cell types throughout the brain and on the blood-brain barrier. In AD, RAGE is up-regulated on cells in the hippocampus, such as astrocytes and microglia.[93] Amyloid is known to bind to these receptors. This may be one way in which the inflammatory cascade is stimulated and thus may lead to cell death (reviewed by Chen *et al.*[94]). Preclinical studies have suggested that blocking this receptor against amyloid binding protects the cell by decreasing plaque formation and inflammation, and it may have an effect on memory functioning (reviewed by Chen *et al.*[94]). Pfizer and the Alzheimer's Disease Cooperative Study are working together on a phase 2 trial of PF-04 494 700, an oral RAGE antagonist. Recently, this dose-finding study has discontinued the highest dose because of evidence of increased behavioral disturbance and worsening cognition (ClinicalTrials.gov: NCT00566397).

Latrepirdine (Dimebon)

After the introduction of memantine, other *N*-methyl-d-aspartate antagonists were sought to expand the field of medications in this class. During one such screen, latrepirdine, once used as a nonselective antihistamine in Russia, was identified. Latrepirdine was studied in animal models of AD[95] and was found to be beneficial in an avoidance conditioning paradigm. The mechanism of action of latrepirdine is unclear because it may also modulate α-amino-3-hydroxy-5-methyl-isoxazole-4-propionic acid and *N*-methyl-d-aspartate receptors and weakly inhibit acheylcholinesterase.[96,97] It also may act by enhancing mitochondrial function.[98] Most recently, in a published trial,[99] 183 patients with mild to moderate AD (an MMSE score of 10–24) were given latrepirdine or placebo. At 26 weeks, the active group had improved performance on the primary cognitive outcome measure, the ADAS-Cog, as well as some secondary outcomes (including behavior, activities of daily living, and global function). In an up to 52-week blinded extension, the latrepirdine group continued to outperform on the ADAS-Cog, and the difference from placebo on the global and activities of daily living measures widened. The most common side effects were dry mouth and depressed mood, each occurring in 14% of the treated group versus 1% to 5% of the placebo group. Other adverse events [hyperhidrosis, insomnia, angina pectoris, asthenia, hyperkalemia, motor dysfunction (repetitive activity), polyuria, atrial flutter, raised blood bilirubin, dyspnea, and musculoskeletal pain] occurred at least twice as frequently in patients given Dimebon in week 26 or 52, but overall frequencies were quite low. Latrepirdine is currently in phase 3 trials in the United States.

Nerve Growth Factor

The loss of acetylcholine neurons is thought to be a significant cause of the memory loss of AD. NGF is a trophic factor for acetylcholine neurons. In a mouse model of Down syndrome, it was seen that increased APP reduced retrograde transport of NGF, and this led to the loss of cholinergic neurons.[100] It is postulated that this may be a mechanism of cholinergic cell loss in AD. NGF has been shown to prevent the death of cholinergic neurons in rats in both aged and lesion models.[101,102] When NGF was implanted in aging rhesus monkeys, it was demonstrated to reverse the loss of cholinergic neurons.[103] In the earliest human trial, NGF was infused into the brain ventricular system of 3 patients, but it caused severe side effects, such as pain from stimulation of dorsal root ganglion nociceptive neurons, Schwann cell migration to the spinal cord, and weight loss.[104] In a search for a more effective way to target NGF at appropriate areas, autologous fibroblasts were genetically modified to express human NGF and then implanted in 8 individuals with AD. There were 2 serious adverse events: the first 2 patients treated, who were sedated but were not under general anesthesia, had intraparenchymal bleeds from sudden movement. The rest of the patients were treated under general anesthesia. In this cohort, there were some statistically significant increases in cerebral metabolism on positron emission tomography scans of several cortical regions at 6 months in 4 participants as opposed to an expected decline with disease progression.[105] More recently, a phase 1 clinical trial looked at NGF in an adenovirus vector. Ten participants at Rush Medical Center and at the University of California, San Diego, were included in an open-label trial looking at various doses of NGF. The researchers continue to follow these patients.[106] Again, there are what appear to be encouraging positron emission tomography findings in 4 patients (ClinicalTrials.gov: NCT00087789). Given the overall tolerability of the neurosurgical procedure under general anesthesia in the phase 1 trial, there is currently a phase 2 clinical trial (ClinicalTrials.gov: NCT00876863) using the highest dose of NGF from the phase 1 trial. The researchers are hoping to enroll 50 subjects to participate in a placebo surgery controlled trial that will look at functional, cognitive, and imaging outcomes over a 2-year period with optional extended follow-up.

Resveratrol

There has been much written about resveratrol, the component found in the skin of red grapes that, in animal studies, has a variety of anti-aging effects, including extending the lifespan of yeast and *Caenorhabditis elegans*.[107,108] In mice, it seemed to improve a variety of aging outcomes (bone health, cholesterol, and coordination), but it did not increase longevity.[109] In an Alzheimer's model of transgenic mice, both red wine[110] and resveratrol[111] significantly reduced plaque formation. Resveratrol is currently being studied in a trial at Mount Sinai with an estimated enrollment of 60 individuals (ClinicalTrials.gov: NCT00678431).

Targeting Tau

Hyperphosphorylated tau protein is the main component of the other neuropathological hallmark of AD, the neurofibrillary tangle. Some scientists argue that the development of neurofibrillary tangles is the earliest neuropathological change in AD and is correlated with disease severity.[112,113] The density of neurofibrillary tangles, particularly in the entorhinal cortex, has been argued to be important in establishing a diagnosis of AD.[114] Mechanistically, hyperphosphorylated tau is also known to interfere with microtubule assembly, which may promote neuronal network breakdown.[115]

Various avenues have been studied in animal models to address tau. One is inhibiting tau kinases,[116] and others involve supporting microtubule assembly.[117,118] Another avenue that has currently passed a phase 2 trial is blocking tau aggregation with methylene blue (MTC).[119] MTC may also enhance mitochondrial function.[120] A 24-week phase 2 trial of 321 patients with mild to moderate AD was blindly continued for up to 84 weeks. At 24 weeks, the patients on active treatment had not declined statistically from the baseline on a standard cognitive measure (the ADAS-Cog). At 50 weeks, there was an 81% reduction in the rate of decline with respect to controls, and the effect size had increased from 24 weeks. Of note, the high test dose was deemed ineffective because of crosslinking to the capsule, and the investigators decided to combine this arm with the placebo arm for statistical purposes. Another significant issue is maintenance of the blinding in this trial because MTC is known to cause the sclera and urine to turn blue.[121] A larger phase 3 trial is planned.

Serotonin Receptors

The discovery of the serotonin 5-hydroxytryptamine-4 (5-HT4) receptor in the past 5 years has provided insights into the signaling pathways and the physiological roles of G protein–coupled receptors in neurons.[55] Animal research has demonstrated the involvement of 5-HT4 receptors in cognitive processes, the protection of neurons via increased secretion of the soluble form of APP, and some evidence of cholinergic stimulation, and all of these are potentially therapeutic in AD.[56] Recent 2-week clinical trials in humans suggest that agonists at the 5-HT4 receptor (eg, PRX-03140) may have a cognitive enhancing effect.[57] This agent was being developed by Epix and was in a phase 2 clinical trial, but after slow enrollment and insufficient financing, the company discontinued this development and left this as an unexplored avenue. Other serotonin receptors have been or are also being explored, such as 5-HT1A (previously described xaliproden) and 5-hydroxytryptamine-6 (SB-742457), which just underwent a phase 1 trial (ClinicalTrials.gov: NCT00551772).

CONCLUSION

This review highlights some of the approaches and directions of research in AD over the past 2 decades. Numerous pathways have been identified to be targeted for potential interventions. There have been leads from serendipitous clinical findings, from observational studies, and from the basic science efforts

Table 1. *Clinical Trials: Duration of recruitment and number of sites.*

Drug	Phase	Length	Country	Organization	Time to Recruit	Reference
Latrepiridine	2/3	6 months (blinded to 1 year), n = 183	Russia	Medivation	5 months, 11 sites	99
LY450139	2	6 weeks, n = 70	United States	Lily	8 months, 6 sites	79
Tarenflurbil	3	18 months, n = 1649	United States	Myriad	18 months, 133 sites	82*
Tramiprosate	3	18 months, n = 1052	United States	Neurochem	13 months, 67 sites	84
AN-1792	2	Terminated, n = 372	United States	Wyeth	15 months 11 sites	85*
Bapineuzumab	2	18 months, n = 240	United States	Wyeth/Elan	7 months, 26 sites	89*
Methylene blue	2	24 weeks (blinded up to 21 months), n = 321	United Kingdom and Singapore	University of Aberdeen and TauRx	Unknown, 17 sites (unregistered trial)	127
Simvastatin (CLASP)	3	18 months, n = 406	United States	ADCS	25 months, 44 sites	38
DHA (AD)	3	18 months, n = 402	United States	ADCS	5 months, 51 sites	61*
DHA (memory complaints)	3	6 months, n = 485	United States	Martek	13 months, 14 sites	61*
Homocysteine	3	18 months, n = 340	United States	ADCS	27 months, 40 sites	73

Abbreviations: AD, Alzheimer's disease; ADCS, Alzheimer's Disease Cooperative Study; CLASP, Cholesterol Lowering Agent to Slow Progression; DHA, docosahexanoic acid.

*Recruitment duration data were gleaned from ClinicalTrials.gov.

Table 2. *Current Medication Trials at Mount Sinai.*

Bapineuzumab
Dimebon
Intravenous immunoglobulin
Nerve growth factor
Receptor for advanced glycation end products inhibitor
Resveritrol
Risperidone discontinuation

NOTE: The general number is 212-241-8329.

at the bench. Regardless of origin, rigorous clinical trials are needed to definitively prove the efficacy of any potential treatment and, importantly, to understand the true risks. In fact, this review summarizes important trials that have been successful, in that they have provided valuable information on potential treatments. Several studies with hopeful agents yielded disappointing results. For example, clinical trials with estrogen, B vitamins, and anti-inflammatory agents identified not only a lack of efficacy but also increased risk for other conditions. Thus, these studies were critically informative and provided directions for avoiding these ineffective treatments.

Over the past 20 years, we have seen the first Food and Drug Administration–approved treatments, which have been shown to have efficacy across the span of the disease, and although the effects are modest, they are consistent and replicable. The ability to answer questions about potential treatments in the next decades will require even more extensive clinical research. Specific requirements for maximizing the efficient development of preventative agents must include the identification of true surrogate markers. The markers must not only provide predictable measures of risk but also be sensitive to interventions that shift risk. Such work cannot be done without large multisite efforts with a commitment to participation in activities such as biomarker databases,[122] longitudinal clinical data registries, and networks of committed subjects and practitioners who are prepared to participate in cutting-edge research. Facilitating drug discovery in AD requires building methods and infrastructure to rapidly conduct large, definitive trials.[123]

The recruitment of subjects to randomized clinical trials remains the great challenge and is considered the major bottleneck to drug development. It is estimated that up to 95% of the delays in study completion are attributable to difficulty in recruiting patients.[124] Table 1 reviews the time and number of sites needed for some of the recent AD trials. Like the experience in cancer trials so recently articulated in the lay press,[125] the regulatory processes and the protection of human subjects require resources, skill, and effort. Clinicians must not be daunted by these requirements and need to accept the challenge of educating patients about the importance of rigorously conducted research. Understanding the value of placebo control trials and the collection of important safety and efficacy information as well as sharing this knowledge with patients is a valuable contribution, particularly when this information can be delivered by trusted practitioners. Patients are eager for the assessment of practitioners about the value of research trials, and many have the altruism to make this contribution to research. A recent report from the Alzheimer's Association[126] indicates that although nearly 40% of patients ask their doctor about research, only 22% of doctors raise the issue with patients and only 21% know of research in their area. This illustrates the significant opportunity that primary care providers have to contribute to the research effort. Table 2 lists the trials currently ongoing at Mount Sinai, and Table 3 provides a list of resources for identifying research in AD both regionally and nationally. Sharing these resources can be one of the most hopeful actions that a clinician can perform for patients.

ACKNOWLEDGMENT

This work was supported by grant AG005138.

DISCLOSURES

Potential conflict of interest: Mary Sano has consulted for the following companies: Novartis, Bristol-Myers Squibb, Medivation, Genentech, Pfizer, and Elan.

Table 3. *Resources to Find Clinical Trials.*

Mount Sinai Alzheimer's Disease Research Center	212-241-8329	http://www.mssm.edu/psychiatry/adrc
Alzheimer's Association: national site	800-272-3900 (24-hour caregiver help line)	http://www.alz.org
Alzheimer's Disease Education and Referral Center	800-438-4380	http://www.nia.nih.gov/alzheimers
ClinicalTrials.gov	Government Web site	http://www.clinicaltrials.gov

REFERENCES

1. Brookmeyer R, Gray S, Kawas C. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. *Am J Public Health* 1998; 88: 1337–1342.
2. Zhu CW, Scarmeas N, Torgan R, et al. Longitudinal study of effects of patient characteristics on direct costs in Alzheimer's disease. *Neurology* 2006; 67: 998–1005.
3. Alzheimer's Association. 2009 Alzheimer's disease facts and figures. *Alzheimers Dement* 2009; 5: 234–270.
4. Deutsch JA, Rocklin KW. Amnesia induced by scopolamine and its temporal variations. *Nature* 1967; 216: 89–90.
5. Fist HS. An evaluation of obstetrical analgesia. *Calif Med* 1954; 80: 91–94.
6. Drachman DA. Memory and cognitive function in man: does the cholinergic system have a specific role? *Neurology* 1977; 27: 783–790.
7. Davis KL, Hollister LE, Overall J, et al. Physostigmine: effects on cognition and effect in normal subjects. *Psychopharmacology (Berl)* 1976; 51: 23–27.
8. Davis KL, Mohs RC. Enhancement of memory processes in Alzheimer's disease with multiple-dose intravenous physostigmine. *Am J Psychiatry* 1982; 139: 1421–1424.
9. Mohs RC, Davis BM, Johns CA, et al. Oral physostigmine treatment of patients with Alzheimer's disease. *Am J Psychiatry* 1985; 142: 28–33.
10. Tariot PN, Farlow MR, Grossberg GT, et al. Memantine treatment in patients with moderate to severe Alzheimer's disease already receiving donepezil: a randomized controlled trial. *JAMA* 2004; 291: 317–324.
11. Reisberg B, Doody R, Stoffler A, et al. Memantine in moderate-to-severe Alzheimer's disease. *N Engl J Med* 2003; 348: 1333–1341.
12. Bauer J, Ganter U, Strauss S, et al. The participation of interleukin-6 in the pathogenesis of Alzheimer's disease. *Res Immunol* 1992; 143: 650–657.
13. Fillit H, Ding WH, Buee L, et al. Elevated circulating tumor necrosis factor levels in Alzheimer's disease. *Neurosci Lett* 1991; 129: 318–320.
14. Griffin WS, Stanley LC, Ling C, et al. Brain interleukin 1 and S-100 immunoreactivity are elevated in Down syndrome and Alzheimer's disease. *Proc Natl Acad Sci U S A* 1989; 86: 7611–7615.
15. Matsubara E, Hirai S, Amari M, et al. Alpha 1-antichymotrypsin as a possible biochemical marker for Alzheimer's-type dementia. *Ann Neurol* 1990; 28: 561–567.
16. McGeer PL, McGeer E, Rogers J, Sibley J. Anti-inflammatory drugs and Alzheimer's disease. *Lancet* 1990; 335: 1037.
17. Hayden KM, Zandi PP, Khachaturian AS, et al. Does NSAID use modify cognitive trajectories in the elderly? The Cache County study. *Neurology* 2007; 69: 275–282.
18. Aisen PS, Davis KL, Berg JD, et al. A randomized controlled trial of prednisone in Alzheimer's disease. Alzheimer's Disease Cooperative Study. *Neurology* 2000; 54: 588–593.
19. Van Gool WA, Weinstein HC, Scheltens P, Walstra GJ. Effect of hydroxychloroquine on progression of dementia in early Alzheimer's disease: an 18-month randomised, double-blind, placebo-controlled study. *Lancet* 2001; 358: 455–460.
20. Aisen PS, Schafer KA, Grundman M, et al. Effects of rofecoxib or naproxen vs placebo on Alzheimer's disease progression: a randomized controlled trial. *JAMA* 2003; 289: 2819–2826.
21. Reines SA, Block GA, Morris JC, et al. Rofecoxib: no effect on Alzheimer's disease in a 1-year, randomized, blinded, controlled study. *Neurology* 2004; 62: 66–71.
22. Breitner JC, Haneuse SJ, Walker R, et al. Risk of dementia and AD with prior exposure to NSAIDs in an elderly community-based cohort. *Neurology* 2009; 72: 1899–1905.
23. Thal LJ, Ferris SH, Kirby L, et al. A randomized, double-blind study of rofecoxib in patients with mild cognitive impairment. *Neuropsychopharmacology* 2005; 30: 1204–1215.
24. Group AR, Lyketsos CG, Breitner JC, et al. Naproxen and celecoxib do not prevent AD in early results from a randomized controlled trial. *Neurology* 2007; 68: 1800–1808.
25. Meinert CL, McCaffrey LD, Breitner JC. Alzheimer's Disease Anti-Inflammatory Prevention Trial: design, methods, and baseline results. *Alzheimers Dement* 2009; 5: 93–104.
26. Chakravarti B, Chakravarti DN. Oxidative modification of proteins: age-related changes. *Gerontology* 2007; 53: 128–139.
27. Nunomura A, Moreira PI, Lee HG, et al. Neuronal death and survival under oxidative stress in Alzheimer's and Parkinson diseases. *CNS Neurol Disord Drug Targets* 2007; 6: 411–423.
28. Sano M, Ernesto C, Thomas RG, et al. A controlled trial of selegiline, alpha-tocopherol, or both as treatment for Alzheimer's disease. The Alzheimer's Disease Cooperative Study. *N Engl J Med* 1997; 336: 1216–1222.
29. Petersen RC, Thomas RG, Grundman M, et al. Vitamin E and donepezil for the treatment of mild cognitive impairment. *N Engl J Med* 2005; 352: 2379–2388.
30. Heart Protection Study Collaborative Group. MRC/BHF heart protection study of antioxidant vitamin supplementation in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002; 360: 23–33.
31. Lippman SM, Klein EA, Goodman PJ, et al. Effect of selenium and vitamin E on risk of prostate cancer and other cancers: the Selenium and Vitamin E Cancer Prevention Trial (SELECT). *JAMA* 2009; 301: 39–51.
32. Wolozin B, Kellman W, Ruosseau P, et al. Decreased prevalence of Alzheimer's disease associated with 3-hydroxy-3-methyglutaryl coenzyme A reductase inhibitors. *Arch Neurol* 2000; 57: 1439–1443.
33. Jick H, Zornberg GL, Jick SS, et al. Statins and the risk of dementia. *Lancet* 2000; 356: 1627–1631.
34. Fassbender K, Simons M, Bergmann C, et al. Simvastatin strongly reduces levels of Alzheimer's disease beta-amyloid peptides Abeta 42 and Abeta 40 in vitro and in vivo. *Proc Natl Acad Sci U S A* 2001; 98: 5856–5861.
35. Sparks DL, Sabbagh MN, Connor DJ, et al. Atorvastatin for the treatment of mild to moderate Alzheimer's disease: preliminary results. *Arch Neurol* 2005; 62: 753–757.

36. Sparks DL, Connor DJ, Sabbagh MN, et al. Circulating cholesterol levels, apolipoprotein E genotype and dementia severity influence the benefit of atorvastatin treatment in Alzheimer's disease: results of the Alzheimer's Disease Cholesterol-Lowering Treatment (ADCLT) trial. *Acta Neurol Scand Suppl* 2006; 185: 3–7.
37. Sjogren M, Gustafsson K, Syversen S, et al. Treatment with simvastatin in patients with Alzheimer's disease lowers both alpha- and beta-cleaved amyloid precursor protein. *Dement Geriatr Cogn Disord* 2003; 16: 25–30.
38. Sano M. Multi-center, randomized, double-blind, placebo-controlled trial of simvastatin to slow the progression of Alzheimer's disease. *Alzheimers Dement* 2008; 4(Suppl 4): T200.
39. Zhou B, Teramukai S, Fukushima M. Prevention and treatment of dementia or Alzheimer's disease by statins: a meta-analysis. *Dement Geriatr Cogn Disord* 2007; 23: 194–201.
40. Heart Protection Study Collaborative Group. MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial. *Lancet* 2002; 360: 7–22.
41. Shepherd J, Blauw GJ, Murphy MB, et al. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet* 2002; 360: 1623–1630.
42. Anonymous Council of the Medical Women's Foundation. An investigation of menopause in one thousand women. *Lancet* 1933; 1: 106–108.
43. Paganini-Hill A, Henderson VW. Estrogen replacement therapy and risk of Alzheimer's disease. *Arch Intern Med* 1996; 156: 2213–2217.
44. Tang MX, Jacobs D, Stern Y, et al. Effect of oestrogen during menopause on risk and age at onset of Alzheimer's disease. *Lancet* 1996; 348: 429–432.
45. Kawas C, Resnick S, Morrison A, et al. A prospective study of estrogen replacement therapy and the risk of developing Alzheimer's disease: the Baltimore Longitudinal Study of Aging. *Neurology* 1997; 48: 1517–1521.
46. Barrett-Connor E, Kritz-Silverstein D. Estrogen replacement therapy and cognitive function in older women. *JAMA* 1993; 269: 2637–2641.
47. Barrett-Connor E, Edelstein S. A prospective study of dehydroepiandrosterone sulfate and cognitive function in an older population: the Rancho Bernardo study. *J Am Geriatr Soc* 1994; 42: 420–423.
48. Brenner DE, Kukull WA, Stergachis A, et al. Postmenopausal estrogen replacement therapy and the risk of Alzheimer's disease: a population-based case-control study. *Am J Epidemiol* 1994; 140: 262–267.
49. DeVoogd T, Nottebohm F. Gonadal hormones induce dendritic growth in the adult avian brain. *Science* 1981; 214: 202–204.
50. Matsumoto A, Arai Y. Neuronal plasticity in the deafferented hypothalamic arcuate nucleus of adult female rats and its enhancement by treatment with estrogen. *J Comp Neurol* 1981; 197: 197–205.
51. Sohrabji F, Miranda RC, Toran-Allerand CD. Estrogen differentially regulates estrogen and nerve growth factor receptor mRNAs in adult sensory neurons. *J Neurosci* 1994; 14: 459–471.
52. Henderson VW, Paganini-Hill A, Miller BL, et al. Estrogen for Alzheimer's disease in women: randomized, double-blind, placebo-controlled trial. *Neurology* 2000; 54: 295–301.
53. Mulnard RA, Cotman CW, Kawas C, et al. Estrogen replacement therapy for treatment of mild to moderate Alzheimer's disease: a randomized controlled trial. Alzheimer's Disease Cooperative Study. *JAMA* 2000; 283: 1007–1015.
54. Wang PN, Liao SQ, Liu RS, et al. Effects of estrogen on cognition, mood, and cerebral blood flow in AD: a controlled study. *Neurology* 2000; 54: 2061–2066.
55. Shumaker SA, Legault C, Kuller L, et al. Conjugated equine estrogens and incidence of probable dementia and mild cognitive impairment in postmenopausal women: Women's Health Initiative Memory Study. *JAMA* 2004; 291: 2947–2958.
56. Craig MC, Maki PM, Murphy DG. The Women's Health Initiative Memory Study: findings and implications for treatment. *Lancet Neurol* 2005; 4: 190–194.
57. Maki PM, Gast MJ, Vieweg AJ, et al. Hormone therapy in menopausal women with cognitive complaints: a randomized, double-blind trial. *Neurology* 2007; 69: 1322–1330.
58. Yaffe K, Krueger K, Cummings SR, et al. Effect of raloxifene on prevention of dementia and cognitive impairment in older women: the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized trial. *Am J Psychiatry* 2005; 162: 683–690.
59. Raloxifene for women with Alzheimer's disease. Available at: http://www.clinicaltrials.gov/ct2/show/NCT00368459. Accessed November 2009.
60. Freund-Levi Y, Eriksdotter-Jonhagen M, Cederholm T, et al. Omega-3 fatty acid treatment in 174 patients with mild to moderate Alzheimer's disease: OmegAD study: a randomized double-blind trial. *Arch Neurol* 2006; 63: 1402–1408.
61. Results from trials of DHA in Alzheimer's disease and age-related cognitive decline. Available at: http://www.alz.org/icad/2010_release_071209_1am_b.asp. Accessed November 2009.
62. van de Rest O, Geleijnse JM, Kok FJ, et al. Effect of fish oil on cognitive performance in older subjects: a randomized, controlled trial. *Neurology* 2008; 71: 430–438.
63. DeKosky ST, Williamson JD, Fitzpatrick AL, et al. Ginkgo biloba for prevention of dementia: a randomized controlled trial. *JAMA* 2008; 300: 2253–2262.
64. Schneider LS. Ginkgo biloba extract and preventing Alzheimer's disease. *JAMA* 2008; 300: 2306–2308.
65. Le Bars PL, Katz MM, Berman N, et al. A placebo-controlled, double-blind, randomized trial of an extract of Ginkgo biloba for dementia. North American EGb Study Group. *JAMA* 1997; 278: 1327–1332.
66. van Dongen MC, van Rossum E, Kessels AG, et al. The efficacy of ginkgo for elderly people with dementia and age-associated memory impairment: new results of a randomized clinical trial. *J Am Geriatr Soc* 2000; 48: 1183–1194.
67. Schneider LS, DeKosky ST, Farlow MR, et al. A randomized, double-blind, placebo-controlled trial of two doses of Ginkgo biloba extract in dementia of the Alzheimer's type. *Curr Alzheimer Res* 2005; 2: 541–551.

68. Birks J, Grimley Evans J. Ginkgo biloba for cognitive impairment and dementia. *Cochrane Database Syst Rev* 2009:; CD003120.
69. Stampfer MJ, Malinow MR, Willett WC, et al. A prospective study of plasma homocyst(e)ine and risk of myocardial infarction in US physicians. *JAMA* 1992; 268: 877–881.
70. Bostom AG, Silbershatz H, Rosenberg IH, et al. Nonfasting plasma total homocysteine levels and all-cause and cardiovascular disease mortality in elderly Framingham men and women. *Arch Intern Med* 1999; 159: 1077–1080.
71. Bostom AG, Rosenberg IH, Silbershatz H, et al. Nonfasting plasma total homocysteine levels and stroke incidence in elderly persons: the Framingham Study. *Ann Intern Med* 1999; 131: 352–355.
72. Seshadri S, Beiser A, Selhub J, et al. Plasma homocysteine as a risk factor for dementia and Alzheimer's disease. *N Engl J Med* 2002; 346: 476–483.
73. Aisen PS, Schneider LS, Sano M, et al. High-dose B vitamin supplementation and cognitive decline in Alzheimer's disease: a randomized controlled trial. *JAMA* 2008; 300: 1774–1783.
74. Viswanathan A, Raj S, Greenberg SM, et al. Plasma Abeta, homocysteine, and cognition: the Vitamin Intervention for Stroke Prevention (VISP) trial. *Neurology* 2009; 72: 268–272.
75. National Horizon Scanning Centre. Xaliproden (Xalprila) for Alzheimer's disease. Available at: www.haps.bham.ac.uk/publichealth/horizon/outputs/documents/2007/april/Xaliproden.pdf. Accessed November 2009.
76. Schechter LE, Smith DL, Rosenzweig-Lipson S, et al. Lecozotan (SRA-333): a selective serotonin 1A receptor antagonist that enhances the stimulated release of glutamate and acetylcholine in the hippocampus and possesses cognitive-enhancing properties. *J Pharmacol Exp Ther* 2005; 314: 1274–1289.
77. Douillet P, Orgogozo JM. What we have learned from the xaliproden Sanofi-Aventis trials. *J Nutr Health Aging* 2009; 13: 365–366.
78. Barten DM, Albright CF. Therapeutic strategies for Alzheimer's disease. *Mol Neurobiol* 2008; 37: 171–186.
79. Siemers ER, Quinn JF, Kaye J, et al. Effects of a gamma-secretase inhibitor in a randomized study of patients with Alzheimer's disease. *Neurology* 2006; 66: 602–604.
80. Wilcock GK, Black SE, Hendrix SB, et al. Efficacy and safety of tarenflurbil in mild to moderate Alzheimer's disease: a randomised phase II trial. *Lancet Neurol* 2008; 7: 483–493.
81. Myriad. Myriad genetics reports results of U.S. phase 3 trial of Flurizan™ in Alzheimer's disease. Flurizan fails to achieve significance on either co-primary endpoint; company has decided to discontinue its development of Flurizan. Available at: http://www.myriad.com/news/release/1170283. Accessed 30 April, 2009.
82. Green RC, Schneider LS, Hendrix S, et al. Safety and efficacy of tarenflurbil in subjects with mild Alzheimer's disease: results from an 18-month multicenter phase 3 trial. *Alzheimers Dement* 2008; 4(Suppl 1): T165.
83. Saumier D, Aisen PS, Gauthier S, et al. Lessons learned in the use of volumetric MRI in therapeutic trials in Alzheimer's disease: the ALZHEMED (tramiprosate) experience. *J Nutr Health Aging* 2009; 13: 370–372.
84. Alzheimer Research Forum. FDA deems U.S. Alzhemed trial results inconclusive. Available at: http://www.alzforum.org/new/detail.asp?id=1647. Accessed November 2009.
85. Gilman S, Koller M, Black RS, et al. Clinical effects of Abeta immunization (AN1792) in patients with AD in an interrupted trial. *Neurology* 2005; 64: 1553–1562.
86. Nicoll JA, Wilkinson D, Holmes C, et al. Neuropathology of human Alzheimer's disease after immunization with amyloid-beta peptide: a case report. *Nat Med* 2003; 9: 448–452.
87. Elan. Elan and Wyeth to initiate phase 3 clinical trial of bapineuzumab (AAB-001) in Alzheimer's disease. Available at: http://www.elan.com/news/2007/20070521.asp.
88. Relkin NR, Szabo P, Adamiak B, et al. 18-month study of intravenous immunoglobulin for treatment of mild Alzheimer's disease. *Neurobiol Aging* 2009; 30: 1728–1736.
89. Wyeth/Elan. Elan and Wyeth announce encouraging top-line results from phase 2 clinical trial of bapineuzumab for Alzheimer's disease. Available at: http://www.wyeth.com/news?nav=display&navTo=/wyeth_html/home/news/pressreleases/2008/1213683456273.html.
90. Wyeth/Elan. Elan and Wyeth plan to amend bapineuzumab phase 3 protocols. Available at: http://www.wyeth.com/news?nav=display&navTo=/wyeth_html/home/news/pressreleases/2009/1238676245463.html.
91. Pfeifer M, Boncristiano S, Bondolfi L, et al. Cerebral hemorrhage after passive anti-Abeta immunotherapy. *Science* 2002; 298: 1379.
92. Vlassara H, Cai W, Crandall J, et al. Inflammatory mediators are induced by dietary glycotoxins, a major risk factor for diabetic angiopathy. *Proc Natl Acad Sci U S A* 2002; 99: 15596–15601.
93. Sasaki N, Toki S, Chowei H, et al. Immunohistochemical distribution of the receptor for advanced glycation end products in neurons and astrocytes in Alzheimer's disease. *Brain Res* 2001; 888: 256–262.
94. Chen X, Walker DG, Schmidt AM, et al. RAGE: a potential target for Abeta-mediated cellular perturbation in Alzheimer's disease. *Curr Mol Med* 2007; 7: 735–742.
95. Lermontova NN, Lukoyanov NV, Serkova TP, et al. Dimebon improves learning in animals with experimental Alzheimer's disease. *Bull Exp Biol Med* 2000; 129: 544–546.
96. Grigorev VV, Dranyi OA, Bachurin SO. Comparative study of action mechanisms of Dimebon and memantine on AMPA- and NMDA-subtypes glutamate receptors in rat cerebral neurons. *Bull Exp Biol Med* 2003; 136: 474–477.
97. Lermontova NN, Redkozubov AE, Shevtsova EF, et al. Dimebon and tacrine inhibit neurotoxic action of beta-amyloid in culture and block L-type Ca(2+) channels. *Bull Exp Biol Med* 2001; 132: 1079–1083.
98. Bachurin SO, Shevtsova EP, Kireeva EG, et al. Mitochondria as a target for neurotoxins and neuroprotective agents. *Ann N Y Acad Sci* 2003; 993: 334–344.
99. Doody RS, Gavrilova SI, Sano M, et al. Effect of Dimebon on cognition, activities of daily living, behaviour, and global function in patients with mild-to-moderate

Alzheimer's disease: a randomised, double-blind, placebo-controlled study. *Lancet* 2008; 372: 207–215.
100. Salehi A, Delcroix JD, Belichenko PV, et al. Increased APP expression in a mouse model of Down's syndrome disrupts NGF transport and causes cholinergic neuron degeneration. *Neuron* 2006; 51: 29–42.
101. Kromer LF. Nerve growth factor treatment after brain injury prevents neuronal death. *Science* 1987; 235: 214–216.
102. Fischer W, Wictorin K, Bjorklund A, et al. Amelioration of cholinergic neuron atrophy and spatial memory impairment in aged rats by nerve growth factor. *Nature* 1987; 329: 65–68.
103. Conner JM, Darracq MA, Roberts J, Tuszynski MH. Nontropic actions of neurotrophins: subcortical nerve growth factor gene delivery reverses age-related degeneration of primate cortical cholinergic innervation. *Proc Natl Acad Sci U S A* 2001; 98: 1941–1946.
104. Eriksdotter Jonhagen M, Nordberg A, Amberla K, et al. Intracerebroventricular infusion of nerve growth factor in three patients with Alzheimer's disease. *Dement Geriatr Cogn Disord* 1998; 9: 246–257.
105. Tuszynski MH, Thal L, Pay M, et al. A phase 1 clinical trial of nerve growth factor gene therapy for Alzheimer's disease. *Nat Med* 2005; 11: 551–555.
106. Alzforum. Drugs in clinical trials: Cere-110. Available at: http://www.alzforum.org/dis/tre/drc/default.asp?type=drugName&drugName=cere. Accessed November 2009.
107. Howitz KT, Bitterman KJ, Cohen HY, et al. Small molecule activators of sirtuins extend *Saccharomyces cerevisiae* lifespan. *Nature* 2003; 425: 191–196.
108. Wood JG, Rogina B, Lavu S, et al. Sirtuin activators mimic caloric restriction and delay ageing in metazoans. *Nature* 2004; 430: 686–689.
109. Pearson KJ, Baur JA, Lewis KN, et al. Resveratrol delays age-related deterioration and mimics transcriptional aspects of dietary restriction without extending life span. *Cell Metab* 2008; 8: 157–168.
110. Wang J, Ho L, Zhao Z, et al. Moderate consumption of Cabernet Sauvignon attenuates Abeta neuropathology in a mouse model of Alzheimer's disease. *FASEB J* 2006; 20: 2313–2320.
111. Karuppagounder SS, Pinto JT, Xu H, et al. Dietary supplementation with resveratrol reduces plaque pathology in a transgenic model of Alzheimer's disease. *Neurochem Int* 2009; 54: 111–118.
112. Terry RD. Tangles precede plaques but don't cause them. *Neurobiol Aging* 2004; 25: 741–742.
113. Haroutunian V, Davies P, Vianna C, et al. Tau protein abnormalities associated with the progression of Alzheimer's disease type dementia. *Neurobiol Aging* 2007; 28: 1–7.
114. Murayama S, Saito Y. Neuropathological diagnostic criteria for Alzheimer's disease. *Neuropathology* 2004; 24: 254–260.
115. Geschwind DH. Tau phosphorylation, tangles, and neurodegeneration: the chicken or the egg? *Neuron* 2003; 40: 457–460.
116. Sun X, Sato S, Murayama O, et al. Lithium inhibits amyloid secretion in COS7 cells transfected with amyloid precursor protein C100. *Neurosci Lett* 2002; 321: 61–64.
117. Michaelis ML, Ansar S, Chen Y, et al. β-Amyloid-induced neurodegeneration and protection by structurally diverse microtubule-stabilizing agents. *J Pharmacol Exp Ther* 2005; 312: 659–668.
118. Butler D, Bendiske J, Michaelis ML, et al. Microtubule-stabilizing agent prevents protein accumulation-induced loss of synaptic markers. *Eur J Pharmacol* 2007; 562: 20–27.
119. Wischik CM, Edwards PC, Lai RY, et al. Selective inhibition of Alzheimer's disease-like tau aggregation by phenothiazines. *Proc Natl Acad Sci U S A* 1996; 93: 11213–11218.
120. Atamna H, Nguyen A, Schultz C, et al. Methylene blue delays cellular senescence and enhances key mitochondrial biochemical pathways. *FASEB J* 2008; 22: 703–712.
121. Gura T. Hope in Alzheimer's fight emerges from unexpected places. *Nat Med* 2008; 14: 894.
122. Cummings JL. Commentary on "a roadmap for the prevention of dementia II: Leon Thal Symposium 2008." Establishing a national biomarker database: utility and incentives. *Alzheimers Dement* 2009; 5: 108–113.
123. Aisen PS. Commentary on "a roadmap for the prevention of dementia II. Leon Thal Symposium 2008." Facilitating Alzheimer's disease drug development in the United States. *Alzheimers Dement* 2009; 5: 125–127.
124. Rowe JC, Elling M, Hazlewood J, Zakhary R. A cure for clinical trials. *McKinsey Q* 2002; 134–141.
125. Kolata G. 40 years war: lack of study volunteers hobbles cancer fight. *New York Times* August 3, 2009.
126. *Physician Study: Final Report*. Chicago, IL: Alzheimer's Association; 2007.
127. 2008 highlights: International Conference on Alzheimer's Disease. Available at: https://www.alz.org/icad/icad_08highlights.asp. Accessed November 2009.

Vaccination as a Therapeutic Approach to Alzheimer's Disease

Thomas Wisniewski, MD,[1–3] and Allal Boutajangout, PhD[1–3]

[1]Department of Neurology, New York University School of Medicine, New York, NY
[2]Department of Pathology, New York University School of Medicine, New York, NY
[3]Department of Psychiatry, New York University School of Medicine, New York, NY

OUTLINE

Pathogenesis of Familial and Sporadic Alzheimer's Disease
Past Human Experience with Active Immunization Targeting AMYLOID BETA
Past Passive Immunization Experience for ad
Phosphorylated tau as an Immune Target
Questions to Address for a New Generation of ad Vaccines
Conclusion

ABSTRACT

Alzheimer's disease is the most common cause of dementia worldwide. Alzheimer's disease is a member of a broad range of neurodegenerative diseases characterized pathologically by the conformational change of a normal protein into a pathological conformer with a high β-sheet content that renders it neurotoxic. In the case of Alzheimer's disease, the normal soluble amyloid β peptide is converted into oligomeric/fibrillar amyloid β. The oligomeric forms of amyloid β have been hypothesized to be the most toxic, whereas fibrillar amyloid β becomes deposited as amyloid plaques and congophilic angiopathy, which both serve as neuropathological markers of the disease. In addition, the accumulation of abnormally phosphorylated tau as soluble toxic oligomers and as neurofibrillary tangles is a critical part of the pathology. Numerous therapeutic interventions are under investigation to prevent and treat Alzheimer's disease. Among the most exciting and advanced of these approaches is vaccination. Immunomodulation is being tried for a range of neurodegenerative disorders, with great success being reported in most model animal trials; however, the much more limited human data have shown more modest clinical success so far, with encephalitis occurring in a minority of patients treated with active immunization. The immunomodulatory approaches for neurodegenerative diseases involve targeting a self-protein, albeit in an abnormal conformation; hence, effective enhanced clearance of the disease-associated conformer has to be balanced with the potential risk of stimulating excessive toxic inflammation within the central nervous system. The design of future immunomodulatory approaches that are more focused is dependent on addressing a number of questions, including when is the best time to start immunization, what are the most appropriate targets for vaccination, and is amyloid central to the pathogenesis of Alzheimer's disease or is it critical to target tau-related pathology also. In this review, we discuss the past experience with vaccination for Alzheimer's disease and the development of possible future strategies that target both amyloid β–related and tau-related pathologies. ***Mt Sinai J Med 77:17–31, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: Alzheimer's disease, amyloid β, immunomodulation, tau, transgenic mice, vaccination.

Alzheimer's disease (AD), affecting approximately 37 million people currently, is the most common cause of dementia worldwide. In the United States, AD is the sixth leading cause of death, with an estimated 5.3 million Americans having AD. By 2050, according to some estimates, 1 of every 85 persons worldwide will be affected by AD.[1] Currently available treatments for AD provide largely symptomatic relief with only minor effects on the course of the disease. There is an urgent need for better therapeutic interventions. Besides

Address Correspondence to:

Thomas Wisniewski
New York University
School of Medicine
New York, NY
Email: thomas.wisniewski@nyumc.org

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20156

immunomodulation, numerous other approaches are being studied, including anti–amyloid β (anti-Aβ) aggregation agents, secretase inhibitors/modulators blocking Aβ production, tau aggregation blockers, agents targeting mitochondria, stem cell therapies, and various neuroprotective strategies.[2] Perhaps the greatest hope for an intervention that shall significantly affect disease progression in the near future comes from the vaccination approaches.[3,4] Certainly in AD transgenic (Tg) mouse models, Aβ-directed immunization has been spectacularly successful with a wide variety of methods. However, significant unanswered questions remain for current and future human trials about the best design of a vaccine, the best target, and the timing of therapy initiation. A key issue that needs to be addressed is the targeting of both Aβ-related and tau-related pathologies.

When the best design for an AD vaccine is being determined, a key issue is the targeting of both Aβ-related and tau-related pathologies.

PATHOGENESIS OF FAMILIAL AND SPORADIC ALZHEIMER'S DISEASE

The pathological hallmarks of AD are the accumulation of Aβ as neuritic plaques and congophilic angiopathy and the accumulation of abnormally phosphorylated tau in the form of neurofibrillary tangles (NFTs). Missense mutations in amyloid precursor protein (APP) or in the presenilin genes [presenilin 1 (PRES1) and PRES2] can cause early-onset, familial forms of AD [familial Alzheimer's disease (FAD)] affecting <4% of AD patients. The most common form of AD is sporadic and of late onset. The dominant theory for the causation of AD has been the amyloid cascade hypothesis.[5,6] This theory currently suggests that the accumulation of Aβ peptides, particularly in a highly toxic oligomeric form, is the primary pathogenic driver that downstream leads to tau hyperphosphorylation, NFT formation, and ultimately synaptic and neuronal loss. Extensive evidence supports this hypothesis in FAD patients and in models of FAD:

1. Inherited forms of AD linked with mutations in the APP gene or in the PRES1 or PRES2 genes are associated with changes in APP processing that favor overproduction of soluble amyloid β (sAβ) or the production of more aggregation prone forms of sAβ such as Aβ 1–42.[7]
2. Down's syndrome, in which there is an extra copy of the APP gene due to trisomy 21, is associated with AD-related pathology at a very early age.[8]
3. In Tg and other models of coexpressed Aβ and tau, Aβ oligomer formation precedes and accentuates tau-related pathology, and this is consistent with the hypothesis that NFT formation is downstream from Aβ aggregation.[9–11]
4. In Tg mouse models of mutant APP overexpression (in which there is no tau pathology), therapeutic prevention and/or the removal of Aβ is associated with cognitive benefits in experimental mice.[12–15]

Importantly, in Tg mouse models of both mutant APP and tau overexpression (with both amyloid-related and tau-related pathologies), prevention of Aβ pathology leads to amelioration of both cognitive deficits and tau-related pathology.[16–18] However, evidence proving that Aβ is central in the common late-onset sporadic form of AD is more limited:

1. A correlation has been shown between biochemically extracted Aβ peptides species from sporadic AD brains with cognitive decline.[19]
2. Isolated Aβ peptide dimers/oligomers from sporadic AD brains have been documented to impair synaptic structure and function.[20]
3. Aβ extracted from sporadic AD patients has been shown to induce amyloid deposits when injected into Tg mice.[21]

A significant problem for the amyloid cascade hypothesis comes from the autopsy data from the initial human active vaccination trial, which is further discussed later. Postmortem analysis was available for 9 subjects in the active immunization arm.[22] All these individuals showed some degree of plaque removal and a reduced Aβ load in comparison with comparable nonimmunized controls. Despite this, there were no differences between placebo and active immunization groups in terms of long-term survival outcome, time to severe dementia, and outcome measures such as the Alzheimer Disease Assessment Scale–Cognitive (ADAS-Cog), Mini Mental State Examination (MMSE), and Disability Assessment for Dementia. This may be related to immunization being begun too late in the disease process; alternatively, one can use these data to suggest that the amyloid cascade hypothesis is an oversimplification. A number of investigators have suggested alternative theories in which the accumulation of Aβ and tau hyperphosphorylation are dual pathways both downstream from a common upstream pathogenic deficit (which remains to be identified).[23–25] In such a scenario, it is essential for immunotherapy to address both of these pathologies to be highly effective. Hence, in this review, we summarize the preclinical and clinical data for both Aβ and phosphorylated tau reduction immunotherapeutic approaches.

> *A number of investigators have suggested [that] the accumulation of Aβ and tau hyperphosphorylation are dual pathways both downstream from a common upstream pathogenic deficit (which remains to be identified).*

PAST HUMAN EXPERIENCE WITH ACTIVE IMMUNIZATION TARGETING AMYLOID BETA

Initial data supporting immunotherapy for AD showed that anti-Aβ antibodies could inhibit Aβ peptide fibrillization/oligomerization and prevent cell culture–based neurotoxicity.[26,27] This led to vaccination of AD Tg mice with Aβ 1–42 or Aβ homologous peptides coinjected with Freund's adjuvant, which demonstrated striking reductions in Aβ deposition and, as a result, elimination of behavioral impairments.[12–15,28,29] Similar effects on Aβ load and behavior have been demonstrated in AD Tg mice by peripheral injections of anti-Aβ monoclonal antibodies, and this indicates that the therapeutic effect of the vaccine is based primarily on eliciting a humoral response.[30,31] In the initial preclinical studies, no toxicity was evident in the treated mice; however, some investigators suggested that the use of nonfibrillogenic, nontoxic Aβ homologous peptides along with approaches that stimulate primarily humoral, T helper 2 (Th2) immunity, in contrast to a primary T helper 1 (Th1) cell–mediated response, might reduce potential toxicity.[32–34] The dramatic biological effect of vaccination in preclinical testing encouraged Elan/Wyeth in April 2000 to launch a randomized, multiple-dose, dose-escalation, double-blind phase I clinical trial with a vaccine designated as AN1792, which contained pre-aggregated Aβ 1–42 and QS21 as an adjuvant. This type of vaccine design was aimed to induce a strong cell-mediated immune response because QS21 is known to be a strong inducer of Th1 lymphocytes.[35] The initial trial was conducted in the United Kingdom and involved 80 patients with mild to moderate AD.[36] This trial was designed to assess the antigenicity and toxicity of multiple-dose immunization with the full-length Aβ 1–42 peptide with QS21. Fifty-three percent of patients developed an anti-Aβ humoral response. During the later stages of the phase I trial, the emulsifier polysorbate 80 was added, causing a greater shift from a Th2-biased response to a proinflammatory Th1 response.[37] In the subsequent phase IIa trial begun in October 2001, 372 patients were enrolled, with 300 receiving the aggregated Aβ 1–42 (AN1792) formulation with QS21 in polysorbate 80. This trial was prematurely terminated in January 2002 when 6% of vaccinated patients manifested symptoms of acute meningoencephalitis (18 of 298 subjects).[35,38,39] Autopsies performed on a limited number of trial patients suggested that striking Aβ clearance of parenchymal plaques had occurred that was similar to what had been reported in the animal studies, and this confirmed the validity of this approach for amyloid clearance in humans.[39–44] In these cases, extensive areas of cerebral cortex were devoid of plaques, with residual plaques having a "moth-eaten" appearance or persisting as "naked" dense cores. This amyloid clearance in most cases was in association with microglia that showed Aβ immunoreactivity, which suggested phagocytosis. Additional striking features were the persistence of amyloid in cerebral vessels and unaltered tau immunoreactive NFTs and neutrophil threads in regions of the cerebral cortex in which plaque clearing had apparently occurred in comparison with regions without clearing.[42–44] Hence, this initial vaccination approach did not address vascular amyloid–related or NFT-related pathology. Some cases also showed a deleterious T cell reaction surrounding some cerebral vessels, which was suggestive of an excessive Th1 immune response. It appeared that the immune reaction triggered by AN1792 was a double-edged sword: the benefits of a humoral response against Aβ were overshadowed in some individuals by a detrimental T cell–mediated inflammatory response.[39,45] The likely involvement of an excess cell-mediated response in mediating toxicity was supported by an analysis of peripheral blood mononuclear cells from trial patients, which were stimulated in vitro with the Aβ peptide, followed by quantification of cytokine secretions by enzyme-linked immunosorbent spot assay.[37] The cells of most responder trial patients mounted interleukin 2 and interferon-γ–positive responses indicative of a class II (CD4+) Th1-type response.[37] Not all patients who received AN1792 responded with antibody production. The majority mounted a humoral response and showed a modest but statistically significant cognitive benefit demonstrated as an improvement on some cognitive testing scales in comparison with the baseline and a slowed rate of disease progression in comparison with the patients who did not form antibodies.[36,46] The follow-up data from the Zurich cohort, a subset of the Elan Wyeth trial,[46,47] indicated that the vaccination approach may be beneficial for human AD patients. In agreement with the findings in the Zurich cohort, immune responders with high antibody titers in the multicenter cohort scored significantly better in composite scores of memory functions in comparison with low responders and nonresponders or with the placebo group of patients.[37] However, it is striking that despite the apparent success in amyloid clearance indicated by the autopsy data, the clinical cognitive benefits were very modest when the active vaccination group was compared to the placebo group.[48] No difference between the antibody responders and the placebo group was found on the ADAS-Cog, Disability Assessment for Dementia, Clinical Dementia Rating scale, MMSE, or Clinical Global Impression of Change. It was only on a 9-item composite neuropsychological test battery that antibody responders had a slight benefit in comparison with the placebo group. These data can be used to suggest that vaccination in this cohort was started too late; hence, tau-related pathology was unaffected by vaccination, and thus the cognitive benefits were small. Alternatively, it can be suggested that the amyloid cascade hypothesis must be an oversimplification of the pathogenesis of sporadic AD. The latter view is supported by a follow-up study of 80 patients in the initial phase I AN1782 trial, of whom 8 came to autopsy.[22] This study showed that despite evidence of very significant amyloid plaque removal in 6 of the 8 autopsy subjects, which correlated with the anti-Aβ titer, there was no evidence of improved survival or an improvement in the time to severe dementia in the overall group.[22]

PAST PASSIVE IMMUNIZATION EXPERIENCE FOR AD

Passive immunization consists of an injection of pre-prepared antibodies to patients, in contrast to active immunization, in which the immune system is stimulated to produce its own antibodies. Passive transfer of exogenous monoclonal anti-Aβ antibodies appears to be the easiest way to fulfill the goal of providing anti-Aβ antibodies without the risk of uncontrolled Th1-mediated autoimmunity. AD Tg model mice treated in this way had significantly reduced Aβ levels and demonstrated cognitive benefits.[30,31] Potential problems with passive immunization include the need for repeated injections in a chronic disease, the high cost, the proper selection of antigen targets, blood-brain barrier penetration, the risk of hemorrhages, and the development of an immune response to the injected antibodies. Several passive immunization trials are underway, with the most advanced being the phase III bapineuzumab trial begun in December 2007.[4] The phase II trial using this anti-Aβ monoclonal antibody was a randomized, double-blind, placebo-controlled trial testing 3 doses in 240 participants. In each of the escalating doses of the antibody, approximately 32 subjects received the active agent, and 28 received placebos. Although the study did not attain statistical significance for the primary efficacy endpoint in the whole study population, in the subgroup of non–apolipoprotein E4 carriers, clinically significant benefits were documented with a number of scales, including the MMSE and Alzheimer's Disease Assessment Scale Battery, over the 18-month trial period. In, addition among non–apolipoprotein E4 carriers, the evaluation of the magnetic resonance imaging results showed less loss of brain volume in treated patients versus control patients. However, it was reported that some patients in the treatment group developed vasogenic edema, a significant adverse reaction. The phase III trial is trying to recruit 800 patients and will run until December 2010.

With a somewhat similar approach, intravenous immunoglobulin (IVIg) is currently in clinical trial for AD, with the rationale being that IVIg contains some anti-Aβ antibodies. In a pilot, open-label study of 8 mild AD patients, IVIg was infused over 6 months, discontinued, and resumed for another 9 months.[49] After each infusion, the plasma Aβ levels increased transiently, with cerebrospinal fluid Aβ being decreased after 6 months. The MMSE increased by an average of 2.5 after 6 months, returned to the baseline after washout, and remained stable with the subsequent IVIg infusions. These promising initial findings clearly need to be repeated in a larger cohort. The attraction of IVIg use is that there is extensive experience using IVIg safely for multiple neurological disorders; however, it is a very expensive treatment, and the percentage of anti-Aβ antibodies in IVIg is extremely low, so this is not likely to be a very specific or highly effective form of treatment.

A particular concern in association with passive immunization is cerebral microhemorrhage. The mechanism of this hemorrhage is thought to be related to Aβ deposition in the form of congophilic amyloid angiopathy (CAA), which causes degeneration of smooth muscle cells and weakening of the blood vessel wall. A number of reports have shown an increase in microhemorrhages in different AD mouse models after passive intraperitoneal immunization with different monoclonal antibodies with a high affinity for $A\beta$ plaques and CAA.[50–52] Microhemorrhages following active immunization in animal models have also been reported but only in 2 studies, so this appears to be less of a problem with this approach.[53,54] In Tg mouse models, $A\beta$ antibodies can both prevent the deposition of vascular amyloid and remove it, thus contributing to vascular repair. On the other hand, the autopsies from the AN1792 trial indicated no clearance of vascular amyloid, and in one of these cases, numerous cortical bleeds were found, which are typically rare in AD patients.[41] This is an important issue as CAA is present in virtually all AD cases, with approximately 20% of AD patients having severe CAA.[55] Furthermore, CAA is present in about 33% of cognitively normal elderly control populations.[56] The need for vascular repair and regeneration during $A\beta$ immunotherapy represents another argument for early treatment as well as an argument favoring subtle clearance over a longer time period.

PHOSPHORYLATED TAU AS AN IMMUNE TARGET

NFTs are a major pathological hallmark of AD. NFTs are intraneuronal inclusion bodies that consist of an accumulation of paired helical filaments (PHFs), which biochemically are mainly composed of abnormally phosphorylated tau. Recently, there has been increasing focus on phosphorylated tau as an immunotherapeutic target.[57–59] In the central nervous system, human tau is expressed in 6 isoforms arising from alternative messenger RNA splicing from a single gene on chromosome 17q21 containing 16 exons (see Figure 1).[60,61] The size range of

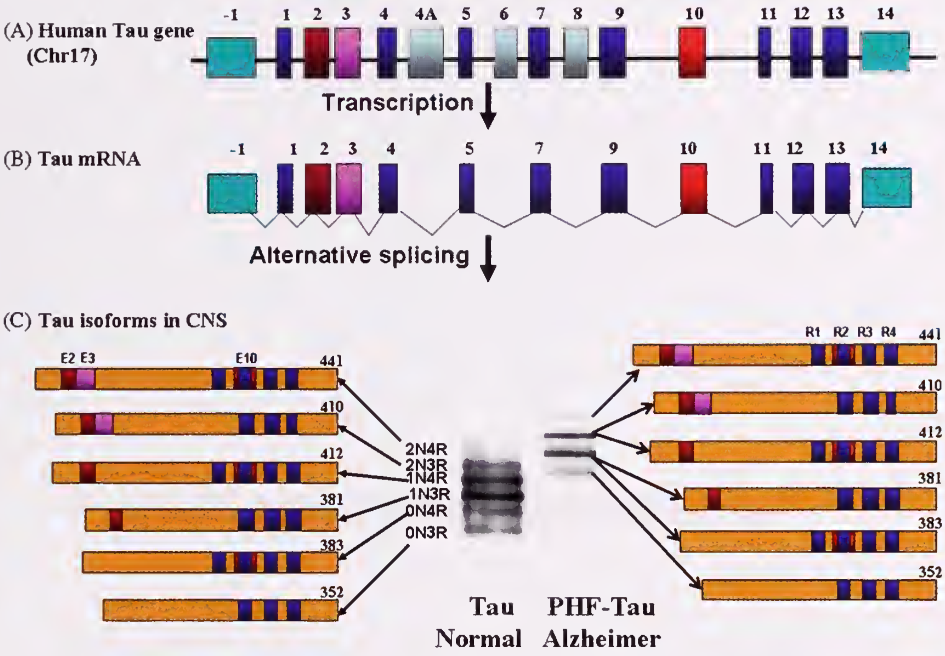


Fig 1. Schematic representation of the human tau gene, which is located on chromosome 17q21 and spans more than 130 kb. This gene is composed of 16 exons. (A) Exons 1 and 14 are transcribed but not translated (turquoise). Exons 4A, 6, and 8 are not transcribed in humans (light blue/charcoal). (B) In the human brain, 6 tau isoforms, ranging from 352 to 441 amino acids, are generated by alternative splicing of exons 2, 3, and 10 (brown/red, pink, and red, respectively) from a single gene. Exons 1, 4, 5, 7, 9, 11, 12, and 13 (blue) are included in all isoforms. Exon 3 is always included with exon 2. The microtubule binding domains are indicated by R1, R2, R3, and R4, which correspond to exons 9, 10, 11, and 12, respectively. (C) The extraction of tau proteins and PHF tau from normal and Alzheimer brains, respectively, shows via immunoblotting 6 bands between 45 and 68 kDa that correspond to different tau isoforms in the normal brain, whereas in PHF tau, 4 bands are detected between 60 and 74 kDa that correspond to the aggregation of 6 hyperphosphorylated tau isoforms in the Alzheimer brain. **Abbreviations:** CNS, central nervous system; mRNA, messenger RNA; PHF, paired helical filament.

the 6 isoforms is 352 to 441 amino acids, which differ by the absence or presence of 29 (exon 2) or 58 (exon 2 + exon 3) amino acid inserts in the amino terminal. The carboxyl terminal half of tau contains 3 or 4 semihomologous repeats of 31 or 32 amino acids encoded by exon 10. The repeats (3R and 4R) correspond to the microtubule binding region of the tau protein (see Figure 1). Stabilization of microtubules by tau is essential for the maintenance of neuronal cell morphology and transport of organelles. In addition, tau has other roles such as interactions with kinesin 1 and the complex dynactin/dynein.[62,63] Tau also plays a crucial role in neuronal cell architecture by interacting with plasma membrane or cytoskeleton proteins such as actin, spectrin, and neurofilament proteins. Several mutations have been described in the tau gene in frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17) and other tauopathies; however no tau gene mutations have been linked to the presence of AD.[64] Most of these mutations affect the binding of tau to microtubules or enhance the aggregation of tau into fibrils. Other intronic mutations that affect the splicing of exon 10 induce an increase in isoforms with 4 repeats. In AD, tau is hyperphosphorylated at all phosphorylated sites, with 9 phosphates per molecule, in comparison with normal brain tau, which has 2 to 3 phosphorylated residues.[65] Other studies have suggested that changes in tau splice forms are related to neurodegeneration. In some animal models expressing mutated tau, there is an increase in 4R tau versus 3R tau.[66] The functional significance of a shift in the 3R tau/4R tau ratio remains unclear, but 4R tau binds microtubules with a higher affinity than 3R tau.[67]

Normal tau and PHF tau differ in the molecular weight and banding pattern, as shown in Figure 1. Normal tau has 6 bands between 45 and 68 kDa, whereas PHF tau has 4 bands between 60 and 74 kDa (see Figure 1).[68,69] The diversity of tau isoforms is related to various posttranslational modifications such as phosphorylation, glycosylation, glycation, ubiquitination, and nitration.[70] The splicing regulation of the tau gene and the relative expression of isoforms are not significantly changed in sporadic AD (Figure 2).[71] Tau has multiple potential phosphorylation sites, which can be detected using specific phosphorylated tau–dependent antibodies (see Figure 3). Seventy-one of the 85 potential phosphorylated sites have been shown to be phosphorylated under physiological or pathological conditions.[72,73] More than 20 protein kinases have been implicated in the phosphorylation of tau proteins, with glycogen synthase kinase-3β and cyclin-independent kinase 5 thought to play the most important roles in phosphorylation under pathological conditions.[72–75]

Several Tg mouse models that express human tau with the FTDP-17 mutation have been produced (see Table 1). Some of these mice display NFTs, neuronal death, and behavioral deficits,[76–85] except for a Tg mouse model that expresses a mutated (N279K) tau that shows behavioral deficits without the formation of NFTs or neuronal loss.[86] In these models, there is disruption of axon transport due to the tau expression. which induces synaptic and neuronal loss. Another Tg tau mouse model was developed that expresses the mutated P301S tau, which shows synaptic loss that precedes tangles formation.[84] The distribution of NFTs in most of these tauopathy models is in contrast to AD because NFTs

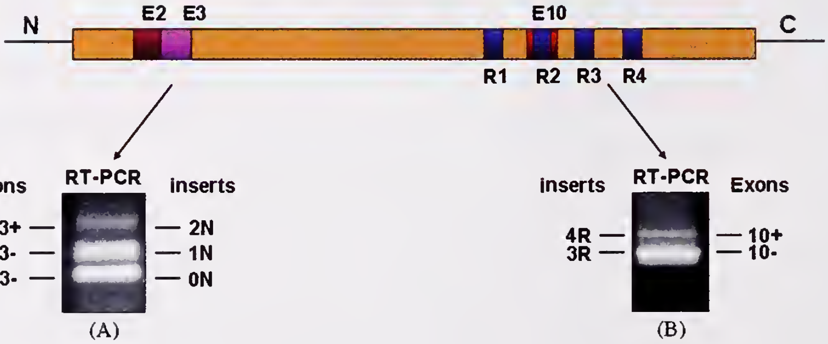


Fig 2. Electrophoresis of RT-PCR amplification products of (A) the 5′ domain of tau mRNAs and (B) the 3′ domain of tau mRNAs. The extraction of RNA was performed from the cerebellar cortex of a sporadic Alzheimer's disease patient. The expression of mRNA by RT-PCR showed different isoforms of human tau detected in the N-terminal (0N, 1N, and 2N) and in the C-terminal (3R and 4R). The plus symbol indicates with an exon, whereas the minus symbol indicates without an exon. **Abbreviations:** mRNA, messenger RNA; RT-PCR, reverse-transcriptase polymerase chain reaction.

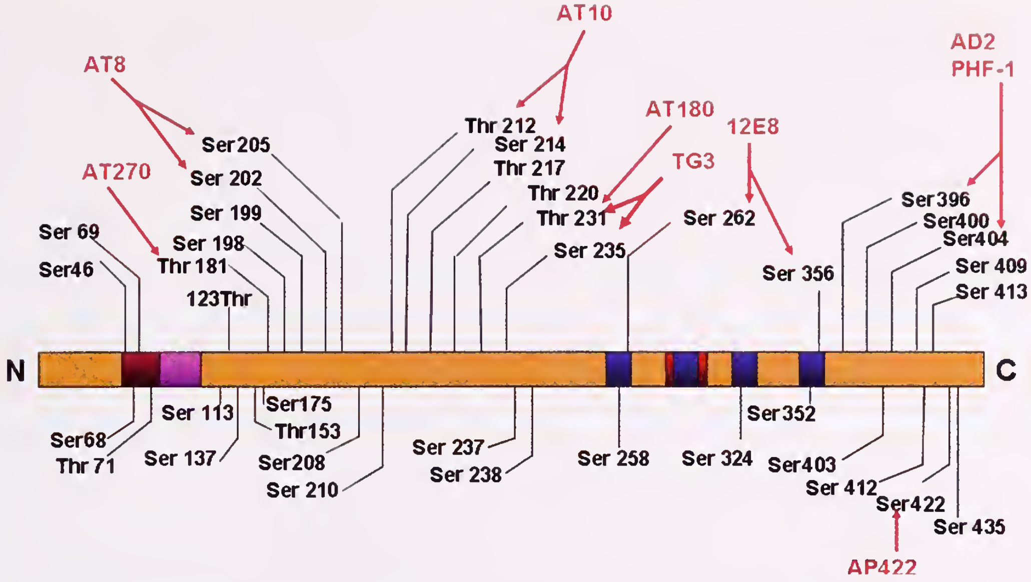


Fig 3. In the human Alzheimer brain, more than 40 phosphorylation sites on tau have been identified and localized in the proline-rich domain and in the C-terminal region. Phosphorylated sites are identified with 8 phosphorylated tau–specific antibodies, as indicated in Figure 1C, with red. It has been suggested that the phosphorylation at Ser 262/356 is responsible for the detachment of tau from microtubules. **Abbreviation:** PHF-1, paired helical filament 1.

are localized in different brain regions such as the brain stem, spinal cord, and frontotemporal cortex instead of the entorhinal region, hippocampus, and neocortex, as observed in AD.[87] In order to generate a more ideal model for AD, other researchers have used a single wild-type human tau to generate a Tg model; however, most of these models did not develop NFT, with the except of 2 models: one expressing ON3R wild-type tau with a few NFTs in aged animals[80] and another with abundant NFTs expressing all 6 human tau isoforms on a knockout background for murine tau.[88,89] The absence of tangles in mice that expressed a single wild-type human tau was likely due to the endogenous tau inhibiting the formation of an NFT-like pathology.

Recently, it has been shown that active immunization of Tg mice (P301L) with a phosphorylated tau peptide (containing the phosphorylated PHF-1 epitopes Ser 396 and Ser 404) for 2 to 5 months could prevent tau-related pathology.[90,91]

Recently, it has been shown that active immunization of Tg mice (P301L) with a phosphorylated tau peptide for 2 to 5 months could prevent tau-related pathology.

These particular phosphorylation epitopes were chosen because these sites have been shown to increase the fibrillogenic nature of tau and contribute to PHF formation.[92,93] Histological and biochemical analyses showed a reduction of aggregated tau in the brain and improved performance on motor tasks.[90] This study clearly documented that it is possible to reduce tau-related pathology with active immunization.

At first examination, it is difficult to understand how an antibody response to a protein that is accumulating intracellularly can have beneficial effects. However, such an outcome is supported by a study of immunization in a Parkinson's disease Tg mouse model with α-synuclein showing a reduction of intracellular α-synuclein aggregates.[94] An additional study has shown that antibodies against Aβ can be internalized in AD neuronal culture models of Aβ accumulation and clear intraneuronal Aβ aggregates via the endosomal-lysosomal pathway.[95] Furthermore, recent evidence has shown that extracellular tau aggregates can be internalized and promote the fibrillization of intracellular full-length tau in a tissue culture model[96] and that the injection of fibrillar tau brain extract into the brains of Tg wild-type expressing mice can induce the formation of human tau into filaments as well as the spread of pathology from the site of injection into neighboring brain regions.[97] This type of infectivity of an abnormal protein conformation from outside the cell has also been demonstrated for polyglutamine aggregates[98] and is well characterized in prion disease.[99,100] Hence, if the spread of PHF

Table 1. *Different Transgenic Models Used to Study Tau-Related Pathology Expressing WT Tau or Mutated Tau Identified in Various FTDP-17 Pedigrees.*

Tau Isoform	Tau Mutation	Presenilin cDNA	APP cDNA	Promoter	NFT	Reference
2N4R	WT	–	–	Thy-1	No	150
2N4R	WT	–	–	Thy-1.2	No	151
2N4R	WT	–	–	Thy-1.2	No	152
0N3R	WT	–	–	HMG-CoAR	No	153
0N3R	WT	–	–	PrP	No	154
3R	WT	–	–	Mouse tubulin Tα1	No	155
0N3R	WT	PS1 M146L	–	HMG-CoAR	No	156
0N3R	WT	PS1 M146L	APP 751 (SL)	HMG-CoAR and Thy-1	No	157
2N4R	WT KoKI	–	–	Thy-1	No	158
6 human isoforms	WT	–	–		Yes	88
0N4R	P301L	–	–	PrP	Yes	159
0N4R	P301L	–	APPsw	PrP	Yes	79
0N4R	P301S	–		Thy-1.2	Yes	160
0N4R	P301L	PS1 M146V	APPsw	Thy-1.2	Yes	10
2N4R	P301L	–	–	Thy-1.2	Yes	9
2N4R	G272V	–	–	PrP	Yes	161
2N4R	M337V	–	–	PDGF	Yes	81
2N4R	R406W	–	–	CaMKII	Yes	162
2N4R	R406W P301L G272V	–	–	Thy-1	Yes	163
2N4R	P301L	–	–	Thy-1	Yes	158
1N2R	G272V P301S			Thy-1.2	Yes	164

Abbreviations: APP, amyloid precursor protein; CaMKII, calcium/calmodulin-dependent protein kinase II; cDNA, complementary DNA; FTDP-17, frontotemporal dementia and parkinsonism linked to chromosome 17; HMG-CoAR, 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase; NFT, neurofibrillary tangle; PDGF, platelet-derived growth factor; PrP, prion protein; PS1, presenilin 1; WT, wild type.

pathology in AD occurs via a prion-like mechanism, anti–phosphorylated tau antibodies would not need to enter cells in order to be effective.

QUESTIONS TO ADDRESS FOR A NEW GENERATION OF AD VACCINES

An initial question that needs to be addressed is when to begin a vaccination protocol. Extensive neuropathological data have established that by the time the earliest clinical signs of AD emerge, Aβ deposition may be close to reaching its peak and that NFT formation and neuronal loss are substantial but have not yet reached peak levels.[101,102] This would suggest that amyloid-directed therapy would need to begin very early, perhaps even before the mild cognitive impairment (MCI) stage, in order to have a maximal effect. This is consistent with the albeit limited autopsy data from the initial AN1792 study, which showed that despite evidence of very significant amyloid plaque removal in 6 of the 8 autopsy subjects, which correlated with the anti-Aβ titer, in the overall group, there was no evidence of improved survival or an improvement in the time to severe dementia.[22] Hence, there is a need for the identification of markers predicting the conversion from normal aging to very mild dementia/MCI. These include cerebrospinal fluid biomarkers such as.[103] Early fluorodeoxyglucose positron emission tomography changes in hippocampal glucose metabolism can predict the conversion of normal cognition to pathologically verified AD.[104] Studies in AD Tg models, using magnetic resonance imaging, suggest that paramagnetic amyloid binding ligands have potential for early amyloid detection and subsequent treatment effects.[105,106] However, direct imaging of

Extensive neuropathological data suggest that amyloid-directed therapy would need to begin very early, perhaps even before the mild cognitive impairment (MCI) stage, in order to have a maximal effect.

amyloid deposits with agents such as Pittsburgh compound B using positron emission tomography is currently the most promising method for the identification of early amyloid deposits and for the identification of patients who will likely convert to MCI from normal aging and from MCI to early AD.[107,108] An alternative approach is to immunize by targeting both Aβ deposition and tau-related pathology. Such an approach has a higher probability of having a clearer effect on the clinical course, even if started when clinical symptoms are evident. Furthermore, if, as discussed previously, tau pathology is not downstream from amyloid deposits but represents a parallel pathology related to a common upstream cause, it will be essential to target tau-related pathology, regardless of how early vaccination treatment is initiated.

An alternative approach is to immunize by targeting both Aβ deposition and tau-related pathology.

Another significant issue that needs to be addressed in future studies is the development of better models for preclinical testing of vaccination approaches. There are many shortcomings to current Tg models of AD pathology. These include the fact that Tg amyloid deposits typically lack the extensive posttranslational modifications of AD amyloid and thus are much more soluble; this presumably allows them to be cleared more easily.[109] The rodent immune system is quite different from the human immune system, and this leads to significant differences in the toxic responses to amyloid deposition.[110] Most current tau Tg models reflect FTDP-related pathology in contrast to AD tau pathology, as discussed previously.[87] Relatively few of the vaccination approaches being developed have been tested in nonhuman primates or other non-Tg models of AD. These may represent more accurate models of the type of immune response that might be elicited in aged humans and better reflect the combination of true human Aβ-related and tau-related pathologies.[111–113] These models include rhesus monkeys, vervet monkeys, mouse lemurs, and aged beagles.[113–118] It is striking that in a recent 22-month active immunization trial of aged beagles, despite an approximately 80% reduction in cortical Aβ immunoreactivity, little cognitive improvement on multiple measures of learning and memory could be detected.[119] However, improvements in some executive functions were found, mirroring the modest improvements seen only in the z score of the Neurological Test Battery among patients in the AN1782 trial. These results reinforce the need to begin immunomodulation very early in the disease progression with a focus on preventing Aβ deposition rather than clearance of preexisting lesions as well as the likely need to target tau-related pathology concurrently.

Active vaccination approaches under development are aiming to avoid the excessive Th1 stimulation that resulted in encephalitis in a proportion of the AN1792 patients. Concurrently, the formulation of any active vaccine also has to overcome the problem of immunosenescence in the target patient population. One promising approach taken by several investigators is to alter the sequence of the Aβ peptide immunogen in order to remove or alter the major Th1 stimulator sites in the carboxyl terminus and the middle portion of Aβ, while focusing on the major Th2 stimulator site in the amino terminus.[28,33,120–122] These Aβ homologous peptide immunogens can be combined with various costimulator epitopes. An example of this approach is a combination with a synthetic, unnatural pan human leukocyte antigen DR–binding epitope (PADRE)[122] or a linkage to viral-like particles[123–125] to induce a primarily humoral immune response. These can be further combined with other immunostimulator carriers. For example, the Aβ Th2 amino terminal epitope can be combined with PADRE and macrophage-derived chemokine in a DNA epitope vaccine to drive robust Th2 responses.[126] The choice of adjuvant is also an important consideration. The use of polysorbate 80, a strong Th1 stimulating adjuvant, in the AN1792 trial is one factor likely contributing to the encephalitis in a minority of patients. The use of adjuvants such as alum that drive primarily a Th2 response is preferable.[29,119] The route of immunization also plays an important role. Stimulating mucosal immunity by nasal, gut, or transcutaneous vaccination has been shown to drive strong Th2 responses.[127–129] An alternative, nonmutually exclusive approach to enhancing vaccine design is to stimulate innate immunity and enable microglia/macrophages to phagocytose amyloid deposits. Over 20 years ago, Wisniewski *et al.*[130] noted that although brain-resident macrophages are unable to phagocytose amyloid, brain-infiltrating macrophages are plaque-competent. A number of recent studies suggest that only a small percentage of plaques are associated with peripheral origin macrophages and that these are required for plaque clearance.[131–133] Vaccination approaches based on this knowledge are now being developed. Stimulation of peripheral macrophages to enter the central nervous system and phagocytose amyloid has been achieved by stimulation of

Toll-like receptor 9 with CpG,[134,135] via blockade of the CD40/CD40L interaction,[136] and by blockade of the transforming growth factor β–Smad2/3 innate signaling pathway.[137] These innate immunity stimulatory approaches can be used alone or in combination with adaptive immunity stimulation. Stimulating the innate immune system has the added potential advantage that it could be effective against both Aβ-related and tau-related pathologies.

Another important issue for future vaccination approaches is the best target for either active or passive immunization. Abundant evidence both in vivo and in vitro suggests that the most toxic species of Aβ are oligomers or Aβ-derived diffusible ligands,[138,139] with a similar line of evidence suggesting that tau oligomers are the most toxic form of phosphorylated tau.[59,84] Active vaccination or the use of monoclonal antibodies that specifically target Aβ oligomers, tau oligomers, or preferably both would be an ideal way to block AD-related toxicity.

> *Active vaccination or the use of monoclonal antibodies that specifically target Aβ oligomers, tau oligomers, or preferably both would be an ideal way to block AD-related toxicity.*

A small number of preclinical studies targeting Aβ oligomers suggest that this methodology is potentially powerful and in need of further development.[140–144] An additional important factor to consider is that emerging evidence suggests that monomeric Aβ peptides have normal physiological functions in the brain such as neuroprotection and modulation of long-term potentiation,[145,146] with phosphorylated tau also having a normal role.[58] Targeting only oligomeric Aβ or tau would avoid the potential of interfering with these physiological functions. A novel immunotherapeutic approach is to target the shared abnormal β-sheet conformation of amyloid proteins with conformationally specific antibodies or active immunization that favors such a conformational response.[140,141,147] Such an approach has the advantage that both Aβ-related and tau-related pathologies would be addressed concurrently.

CONCLUSION

Numerous therapeutic approaches are under development for AD; however, harnessing the immune system to clear both Aβ-related and tau-related pathologies is perhaps the most promising and advanced approach. Abnormal protein conformation is thought to be the underlying pathogenesis of not only AD but also a long list of neurodegenerative conditions, such as prion disease, Parkinson's disease, and Huntington's chorea, with immunomodulation having the potential to be a disease-altering therapeutic approach for all these disorders. For example, it has been shown that prion-directed mucosal vaccination can prevent infection from an exogenous source.[148,149] Ultimately, directing the immune system to clear the highly toxic, abnormal oligomeric conformers that characterize multiple neurodegenerative diseases has the potential to dramatically alter the course of a wide spectrum of age-associated diseases.

ACKNOWLEDGMENT

This article was supported by grants AG20245 and AG15408 from the National Institutes of Health and by grant IIRG-06-26434 from the Alzheimer's Association.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM. Forecasting the global burden of Alzheimer's disease. *Alzheimers Dement* 2007; 3: 186–191.
2. Rafii MS, Aisen PS. Recent developments in Alzheimer's disease therapeutics. *BMC Med* 2009; 7: 7.
3. Brody DL, Holtzman DM. Active and passive immunotherapy for neurodegenerative diseases. *Annu Rev Neurosci* 2008; 31: 175–193.
4. Wisniewski T, Konietzko U. Amyloid-β immunization for Alzheimer's disease. *Lancet Neurol* 2008; 7: 805–811.
5. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* 2002; 297: 353–356.
6. Tanzi RE, Bertram L. Twenty years of the Alzheimer's disease amyloid hypothesis: a genetic perspective. *Cell* 2005; 120: 545–555.
7. Hardy J. A hundred years of Alzheimer's disease research. *Neuron* 2006; 52: 3–13.
8. Lemere CA, Blusztajn JK, Yamaguchi H, et al. Sequence of deposition of heterogeneous amyloid β-peptides and APO E in Down syndrome: implications for initial events in amyloid plaque formation. *Neurobiol Dis* 1996; 3: 16–32.
9. Gotz J, Chen F, van DJ, Nitsch RM. Formation of neurofibrillary tangles in P301l tau transgenic mice

induced by Abeta 42 fibrils. *Science* 2001; 293: 1491–1495.

10. Oddo S, Caccamo A, Shepherd JD, et al. Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular Abeta and synaptic dysfunction. *Neuron* 2003; 39: 409–421.
11. King ME, Kan HM, Baas PW, et al. Tau-dependent microtubule disassembly initiated by prefibrillar beta-amyloid. *J Cell Biol* 2006; 175: 541–546.
12. Schenk D, Barbour R, Dunn W, et al. Immunization with amyloid-β attenuates Alzheimer disease-like pathology in the PDAPP mice. *Nature* 1999; 400: 173–177.
13. Sigurdsson EM, Scholtzova H, Mehta P, et al. Immunization with a non-toxic/non-fibrillar amyloid-β homologous peptide reduces Alzheimer's disease associated pathology in transgenic mice. *Am J Pathol* 2001; 159: 439–447.
14. Morgan D, Diamond DM, Gottschall PE, et al. Aβ peptide vaccination prevents memory loss in an animal model of Alzheimer's disease. *Nature* 2001; 408: 982–985.
15. Janus C, Pearson J, McLaurin J, et al. Aβ peptide immunization reduces behavioural impairment and plaques in a model of Alzheimer's disease. *Nature* 2000; 408: 979–982.
16. Oddo S, Caccamo A, Tran L, et al. Temporal profile of amyloid-beta (Abeta) oligomerization in an in vivo model of Alzheimer disease. A link between Abeta and tau pathology. *J Biol Chem* 2006; 281: 1599–1604.
17. Blurton-Jones M, LaFerla FM. Pathways by which Abeta facilitates tau pathology. *Curr Alzheimer Res* 2006; 3: 437–448.
18. McKee AC, Carreras I, Hossain L, et al. Ibuprofen reduces Abeta, hyperphosphorylated tau and memory deficits in Alzheimer mice. *Brain Res* 2008; 1207: 225–236.
19. Naslund J, Haroutunian V, Mohs R, et al. Correlation between elevated levels of amyloid beta-peptide in the brain and cognitive decline. *JAMA* 2000; 283: 1571–1577.
20. Shankar GM, Li S, Mehta TH, et al. Amyloid-beta protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory. *Nat Med* 2008; 14: 837–842.
21. Meyer-Luehmann M, Coomaraswamy J, Bolmont T, et al. Exogenous induction of cerebral beta-amyloidogenesis is governed by agent and host. *Science* 2006; 313: 1781–1784.
22. Holmes C, Boche D, Wilkinson D, et al. Long term effects of Aβ42 immunization in Alzheimer's disease: immune response, plaque removal and clinical function. *Lancet* 2008; 372: 216–223.
23. Small SA, Duff K. Linking Abeta and tau in late-onset Alzheimer's disease: a dual pathway hypothesis. *Neuron* 2008; 60: 534–542.
24. Castellani RJ, Lee HG, Zhu X, et al. Alzheimer disease pathology as a host response. *J Neuropathol Exp Neurol* 2008; 67: 523–531.
25. Shioi J, Georgakopoulos A, Mehta P, et al. FAD mutants unable to increase neurotoxic Abeta 42 suggest that mutation effects on neurodegeneration may be independent of effects on Abeta. *J Neurochem* 2007; 101: 674–681.
26. Solomon B, Koppel R, Frankel D, Hanan-Aharon E. Disaggregation of Alzheimer β-amyloid by site-directed mAb. *Proc Natl Acad Sci U S A* 1997; 94: 4109–4112.
27. Solomon B. Antibody-mediated immunotherapy for Alzheimer's disease. *Curr Opin Investig Drugs* 2007; 8: 519–524.
28. Sigurdsson EM, Knudsen EL, Asuni A, et al. An attenuated immune response is sufficient to enhance cognition in an Alzheimer's disease mouse model immunized with amyloid-β derivatives. *J Neurosci* 2004; 24: 6277–6282.
29. Asuni A, Boutajangout A, Scholtzova H, et al. Aβ derivative vaccination in alum adjuvant prevents amyloid deposition and does not cause brain microhemorrhages in Alzheimer's model mice. *Eur J Neurosci* 2006; 24: 2530–2542.
30. Bard F, Cannon C, Barbour R, et al. Peripherally administered antibodies against amyloid beta-peptide enter the central nervous system and reduce pathology in a mouse model of Alzheimer disease. *Nat Med* 2000; 6: 916–919.
31. DeMattos RB, Bales KR, Cummins DJ, et al. Peripheral anti-A beta antibody alters CNS and plasma A beta clearance and decreases brain A beta burden in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci U S A* 2001; 98: 8850–8855.
32. Lemere CA, Maron R, Selkoe DJ, Weiner HL. Nasal vaccination with beta-amyloid peptide for the treatment of Alzheimer's disease. *DNA Cell Biol* 2001; 20: 705–711.
33. Sigurdsson EM, Scholtzova H, Mehta P, et al. Immunization with a nontoxic/nonfibrillar amyloid-β homologous peptide reduces Alzheimer's disease associated pathology in transgenic mice. *Am J Pathol* 2001; 159: 439–447.
34. Sigurdsson EM, Frangione B, Wisniewski T. Immunization for Alzheimer's disease. *Drug Dev Res* 2002; 56: 135–142.
35. Wisniewski T, Frangione B. Immunological and anti-chaperone therapeutic approaches for Alzheimer's disease. *Brain Pathol* 2005; 15: 72–77.
36. Bayer AJ, Bullock R, Jones RW, et al. Evaluation of the safety and immunogenicity of synthetic Aβ 42 (AN1792) in patients with AD. *Neurology* 2005; 64: 94–101.
37. Pride M, Seubert P, Grundman M, et al. Progress in the active immunotherapeutic approach to Alzheimer's disease: clinical investigations into AN1792-associated meningoencephalitis. *Neurodegener Dis* 2008; 5: 194–196.
38. Wisniewski T. Practice point commentary on "Clinical effects of Aβ immunization (AN1792) in patients with AD in an interrupted trial." *Nat Clin Pract Neurol* 2005; 1: 84–85.
39. Boche D, Nicoll JA. The role of the immune system in clearance of Abeta from the brain. *Brain Pathol* 2008; 18: 267–278.
40. Bombois S, Maurage CA, Gompel M, et al. Absence of beta-amyloid deposits after immunization in Alzheimer disease with Lewy body dementia. *Arch Neurol* 2007; 64: 583–587.
41. Ferrer I, Boada RM, Sanchez Guerra ML, et al. Neuropathology and pathogenesis of encephalitis following amyloid-beta immunization in Alzheimer's disease. *Brain Pathol* 2004; 14: 11–20.

42. Masliah E, Hansen L, Adame A, et al. Aβ vaccination effects on plaque pathology in the absence of encephalitis in Alzheimer disease. *Neurology* 2005; 64: 129–131.
43. Nicoll JA, Wilkinson D, Holmes C, et al. Neuropathology of human Alzheimer disease after immunization with amyloid-beta peptide: a case report. *Nat Med* 2003; 9: 448–452.
44. Nicoll JA, Barton E, Boche D, et al. Abeta species removal after abeta42 immunization. *J Neuropathol Exp Neurol* 2006; 65: 1040–1048.
45. Sadowski M, Wisniewski T. Disease modifying approaches for Alzheimer's pathology. *Curr Pharm Des* 2007; 13: 1943–1954.
46. Hock C, Konietzko U, Straffer JR, et al. Antibodies against β-amyloid slow cognitive decline in Alzheimer's disease. *Neuron* 2003; 38: 547–554.
47. Hock C, Konietzko U, Paspassotiropoulos A, et al. Generation of antibodies specific for β-amyloid by vaccination of patients with Alzheimer disease. *Nat Med* 2002; 8: 1270–1276.
48. Gilman S, Koller M, Black RS, et al. Clinical effects of Aβ immunization (AN1792) in patients with AD in an interrupted trial. *Neurology* 2005; 64: 1553–1562.
49. Relkin NR, Szabo P, Adamiak B, et al. 18-month study of intravenous immunoglobulin for treatment of mild Alzheimer disease. *Neurobiol Aging* 2009; 30: 1728–1736.
50. Pfeifer M, Boncristiano S, Bondolfi L, et al. Cerebral hemorrhage after passive anti-Aβ immunotherapy. *Science* 2002; 298: 1379.
51. Wilcock DM, Rojiani A, Rosenthal A, et al. Passive immunization against Abeta in aged APP-transgenic mice reverses cognitive deficits and depletes parenchymal amyloid deposits in spite of increased vascular amyloid and microhemorrhage. *J Neuroinflamm* 2004; 1: 24.
52. Racke MM, Boone LI, Hepburn DL, et al. Exacerbation of cerebral amyloid angiopathy-associated microhemorrhages in amyloid precursor protein transgenic mice by immunotherapy is dependent on antibody recognition of deposited forms of amyloid beta. *J Neurosci* 2005; 25: 629–636.
53. Wilcock DM, Jantzen PT, Li Q, et al. Amyloid-beta vaccination, but not nitro-nonsteroidal anti-inflammatory drug treatment, increases vascular amyloid and microhemorrhage while both reduce parenchymal amyloid. *Neuroscience* 2007; 144: 950–960.
54. Petrushina I, Ghochikyan A, Mkrtichyan M, et al. Mannan-Abeta28 conjugate prevents Abeta-plaque deposition, but increases microhemorrhages in the brains of vaccinated Tg2576 (APPsw) mice. *J Neuroinflamm* 2008; 5: 42.
55. Jellinger KA. Alzheimer disease and cerebrovascular pathology: an update. *J Neural Transm* 2002; 109: 813–836.
56. Zhang-Nunes SX, Maat-Schieman ML, Van Duinen SG, et al. The cerebral beta-amyloid angiopathies: hereditary and sporadic. *Brain Pathol* 2006; 16: 30–39.
57. Sigurdsson EM. Immunotherapy targeting pathological tau protein in Alzheimer's disease and related tauopathies. *J Alzheimers Dis* 2008; 15: 157–168.
58. Noble W, Garwood CJ, Hanger DP. Minocycline as a potential therapeutic agent in neurodegenerative disorders characterised by protein misfolding. *Prion* 2009; 3: 78–83.
59. Kayed R, Jackson GR. Prefilament tau species as potential targets for immunotherapy for Alzheimer disease and related disorders. *Curr Opin Immunol* 2009; 21: 359–363.
60. Goedert M, Wischik CM, Crowther RA, et al. Cloning and sequencing of the cDNA encoding a core protein of the paired helical filament of Alzheimer disease: identification as the microtubule-associated protein tau. *Proc Natl Acad Sci U S A* 1988; 85: 4051–4055.
61. Goedert M, Spillantini MG, Jakes R, et al. Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron* 1989; 3: 519–526.
62. Utton MA, Noble WJ, Hill JE, et al. Molecular motors implicated in the axonal transport of tau and alpha-synuclein. *J Cell Sci* 2005; 118: 4645–4654.
63. Magnani E, Fan J, Gasparini L, et al. Interaction of tau protein with the dynactin complex. *EMBO J* 2007; 26: 4546–4554.
64. Gendron TF, Petrucelli L. The role of tau in neurodegeneration. *Mol Neurodegener* 2009; 4: 13.
65. Morishima-Kawashima M, Hasegawa M, Takio K, et al. Hyperphosphorylation of tau in PHF. *Neurobiol Aging* 1995; 16: 365–371.
66. Sergeant N, David JP, Lefranc D, et al. Different distribution of phosphorylated tau protein isoforms in Alzheimer's and Pick's diseases. *FEBS Lett* 1997; 412: 578–582.
67. Butner KA, Kirschner MW. Tau protein binds to microtubules through a flexible array of distributed weak sites. *J Cell Biol* 1991; 115: 717–730.
68. Brion JP, Flament-Durand J, Dustin P. Alzheimer's disease and tau proteins. *Lancet* 1986; 2: 1098.
69. Brion JP. Immunological demonstration of tau protein in neurofibrillary tangles of Alzheimer's disease. *J Alzheimers Dis* 2006; 9: 177–185.
70. Wang JZ, Liu F. Microtubule-associated protein tau in development, degeneration and protection of neurons. *Prog Neurobiol* 2008; 85: 148–175.
71. Boutajangout A, Boom A, Leroy K, Brion JP. Expression of tau mRNA and soluble tau isoforms in affected and non-affected brain areas in Alzheimer's disease. *FEBS Lett* 2004; 576: 183–189.
72. Buee L, Bussiere T, Buee-Scherrer V, et al. Tau protein isoforms, phosphorylation and role in neurodegenerative disorders. *Brain Res Brain Res Rev* 2000; 33: 95–130.
73. Sergeant N, Bretteville A, Hamdane M, et al. Biochemistry of tau in Alzheimer's disease and related neurological disorders. *Expert Rev Proteomics* 2008; 5: 207–224.
74. Baumann K, Mandelkow EM, Biernat J, et al. Abnormal Alzheimer-like phosphorylation of tau-protein by cyclin-dependent kinases cdk2 and cdk5. *FEBS Lett* 1993; 336: 417–424.
75. Hamdane M, Sambo AV, Delobel P, et al. Mitotic-like tau phosphorylation by p25-Cdk5 kinase complex. *J Biol Chem* 2003; 278: 34026–34034.
76. Gotz J, Barmettler R, Ferrari A, et al. In vivo analysis of wild-type and FTDP-17 tau transgenic mice. *Ann N Y Acad Sci* 2000; 920: 126–133.
77. Gotz J, Chen F, Barmettler R, Nitsch RM. Tau filament formation in transgenic mice expressing P301L tau. *J Biol Chem* 2001; 276: 529–534.

78. Gotz J, Ittner LM. Animal models of Alzheimer's disease and frontotemporal dementia. *Nat Rev Neurosci* 2008; 9: 532–544.
79. Lewis J, Dickson D, Lin WL, et al. Enhanced neurofibrillary degeneration in transgenic mice expressing mutant tau and APP. *Science* 2001; 293: 1487–1491.
80. Ishihara T, Zhang B, Higuchi K, et al. Age-dependent induction of congophilic neurofibrillary tau inclusions in tau transgenic mice. *Am J Pathol* 2001; 158: 555–562.
81. Tanemura K, Akagi T, Murayama M, et al. Formation of filamentous tau aggregations in transgenic mice expressing V337M human tau. *Neurobiol Dis* 2001; 8: 1036–1045.
82. Egashira N, Iwasaki K, Takashima A, et al. Altered depression-related behavior and neurochemical changes in serotonergic neurons in mutant R406W human tau transgenic mice. *Brain Res* 2005; 1059: 7–12.
83. Lee VM, Kenyon TK, Trojanowski JQ. Transgenic animal models of tauopathies. *Biochim Biophys Acta* 2005; 1739: 251–259.
84. Yoshiyama Y, Higuchi M, Zhang B, et al. Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model. *Neuron* 2007; 53: 337–351.
85. Murakami T, Paitel E, Kawarabayashi T, et al. Cortical neuronal and glial pathology in TgTauP301L transgenic mice: neuronal degeneration, memory disturbance, and phenotypic variation. *Am J Pathol* 2006; 169: 1365–1375.
86. Taniguchi T, Doe N, Matsuyama S, et al. Transgenic mice expressing mutant (N279K) human tau show mutation dependent cognitive deficits without neurofibrillary tangle formation. *FEBS Lett* 2005; 579: 5704–5712.
87. Zilka N, Korenova M, Novak M. Misfolded tau protein and disease modifying pathways in transgenic rodent models of human tauopathies. *Acta Neuropathol* 2009; 118: 71–86.
88. Andorfer C, Kress Y, Espinoza M, et al. Hyperphosphorylation and aggregation of tau in mice expressing normal human tau isoforms. *J Neurochem* 2003; 86: 582–590.
89. Andorfer C, Acker CM, Kress Y, et al. Cell-cycle reentry and cell death in transgenic mice expressing nonmutant human tau isoforms. *J Neurosci* 2005; 25: 5446–5454.
90. Asuni AA, Boutajangout A, Quartermain D, Sigurdsson EM. Immunotherapy targeting pathological tau conformers in a tangle mouse model reduces brain pathology with associated functional improvements. *J Neurosci* 2007; 27: 9115–9129.
91. Sigurdsson EM. Tau-focused immunotherapy for Alzheimer's disease and related tauopathies. *Curr Alzheimer Res* 2009; 6: 446–450.
92. Eidenmuller J, Fath T, Maas T, et al. Phosphorylation-mimicking glutamate clusters in the proline-rich region are sufficient to simulate the functional deficiencies of hyperphosphorylated tau protein. *Biochem J* 2001; 357: 759–767.
93. Fath T, Eidenmuller J, Brandt R. Tau-mediated cytotoxicity in a pseudohyperphosphorylation model of Alzheimer's disease. *J Neurosci* 2002; 22: 9733–9741.
94. Masliah E, Rockenstein E, Adame A, et al. Effects of alpha-synuclein immunization in a mouse model of Parkinson's disease. *Neuron* 2005; 46: 857–868.
95. Tampellini D, Magrane J, Takahashi RH, et al. Internalized antibodies to the Abeta domain of APP reduce neuronal Abeta and protect against synaptic alterations. *J Biol Chem* 2007; 282: 18895–18906.
96. Frost B, Jacks RL, Diamond MI. Propagation of tau misfolding from the outside to the inside of a cell. *J Biol Chem* 2009; 284: 12845–12852.
97. Clavaguera F, Bolmont T, Crowther RA, et al. Transmission and spreading of tauopathy in transgenic mouse brain. *Nat Cell Biol* 2009; 11: 909–913.
98. Ren PH, Lauckner JE, Kachirskaia I, et al. Cytoplasmic penetration and persistent infection of mammalian cells by polyglutamine aggregates. *Nat Cell Biol* 2009; 11: 219–225.
99. Aguzzi A. Cell biology: beyond the prion principle. *Nature* 2009; 459: 924–925.
100. Sadowski M, Verma A, Wisniewski T. Infectious disease of the nervous system: prion diseases. In: Bradley WG, Daroff RB, Fenichel GM, Jankovic J, eds. *Neurology in Clinical Practice*. Philadelphia, PA: Elsevier; 2008:; 1567–1581.
101. Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. *Ann Neurol* 1999; 45: 358–368.
102. Tarawneh R, Holtzman DM. Critical issues for successful immunotherapy in Alzheimer's disease: development of biomarkers and methods for early detection and intervention. *CNS Neurol Disord Drug Targets* 2009; 8: 144–159.
103. Fagan AM, Roe CM, Xiong C, et al. Cerebrospinal fluid tau/beta-amyloid (42) ratio as a prediction of cognitive decline in nondemented older adults. *Arch Neurol* 2007; 64: 343–349.
104. Mosconi L, Mistur R, Switalski R, et al. FDG-PET changes in brain glucose metabolism from normal cognition to pathologically verified Alzheimer's disease. *Eur J Nucl Med Mol Imaging* 2009; 36: 811–822.
105. Sigurdsson EM, Wadghiri YZ, Mosconi L, et al. A non-toxic ligand for voxel-based MRI analysis of plaques in AD transgenic mice. *Neurobiol Aging* 2008; 29: 836–877.
106. Scholtzova H, Wadghiri YZ, Douadi M, et al. A NMDA receptor antagonist leads to behavioral improvement and amyloid reduction in Alzheimer's disease model transgenic mice shown by micro-magnetic resonance imaging. *J Neurosci Res* 2008; 86: 2784–2791.
107. Rowe CC, Ng S, Ackermann U, et al. Imaging beta-amyloid burden in aging and dementia. *Neurology* 2007; 68: 1718–1725.
108. Forsberg A, Engler H, Almkvist O, et al. PET imaging of amyloid deposition in patients with mild cognitive impairment. *Neurobiol Aging* 2008; 29: 1456–1465.
109. Roher AE, Kokjohn TA. Of mice and men: the relevance of transgenic mice Abeta immunizations to Alzheimer's disease. *J Alzheimers Dis* 2002; 4: 431–434.
110. Tenner AJ, Fonseca MI. The double-edged flower: roles of complement protein C1q in neurodegenerative diseases. *Adv Exp Med Biol* 2006; 586: 153–176.
111. Gandy S, DeMattos RB, Lemere CA, et al. Alzheimer's Aβ vaccination of rhesus monkeys (*Macaca mulatta*). *Alzheimer Dis Assoc Disord* 2004; 18: 44–46.
112. Lemere CA, Beierschmitt A, Iglesias M, et al. Alzheimer's disease abeta vaccine reduces central nervous system abeta levels in a non-human primate, the Caribbean vervet. *Am J Pathol* 2004; 165: 283–297.

113. Trouche SG, Asuni A, Rouland S, et al. Antibody response and plasma Aβ 1–40 in young *Microcebus murinus* primates immunized with Aβ 1–42 and its derivatives. *Vaccine* 2009; 27: 957–964.
114. Gandy S, DeMattos RB, Lemere CA, et al. Alzheimer's Aβ vaccination of rhesus monkeys (*Macaca mulatta*). *Mech Ageing Dev* 2004; 125: 149–151.
115. Lemere CA, Iglesias M, Spooner ET, et al. Aβ immunization in aged vervet monkeys reduces Aβ levels in brain and CSF. *Soc Neurosci Abstr* 2003; 133: 5.
116. Lemere CA, Maier M, Jiang L, et al. Amyloid-beta immunotherapy for the prevention and treatment of Alzheimer disease: lessons from mice, monkeys, and humans. *Rejuvenation Res* 2006; 9: 77–84.
117. Vasilevko V, Head E. Immunotherapy in a natural model of Abeta pathogenesis: the aging beagle. *CNS Neurol Disord Drug Targets* 2009; 8: 98–113.
118. Cotman CW, Head E. The canine (dog) model of human aging and disease: dietary, environmental and immunotherapy approaches. *J Alzheimers Dis* 2008; 15: 685–707.
119. Head E, Pop V, Vasilevko V, et al. A two-year study with fibrillar beta-amyloid (Abeta) immunization in aged canines: effects on cognitive function and brain Abeta. *J Neurosci* 2008; 28: 3555–3566.
120. Cribbs DH, Ghochikyan A, Vasilevko V, et al. Adjuvant-dependent modulation of Th1 and Th2 responses to immunization with beta-amyloid. *Int Immunol* 2003; 15: 505–514.
121. Maier M, Seabrook TJ, Lazo ND, et al. Short amyloid-beta (Abeta) immunogens reduce cerebral Abeta load and learning deficits in an Alzheimer's disease mouse model in the absence of an Abeta-specific cellular immune response. *J Neurosci* 2006; 26: 4717–4728.
122. Agadjanyan MG, Ghochikyan A, Petrushina I, et al. Prototype Alzheimer's disease vaccine using the immunodominant B cell epitope from beta-amyloid and promiscuous T cell epitope pan HLA DR-binding peptide. *J Immunol* 2005; 174: 1580–1586.
123. Zamora E, Handisurya A, Shafti-Keramat S, et al. Papillomavirus-like particles are an effective platform for amyloid-beta immunization in rabbits and transgenic mice. *J Immunol* 2006; 177: 2662–2670.
124. Chackerian B, Rangel M, Hunter Z, Peabody DS. Virus and virus-like particle-based immunogens for Alzheimer's disease induce antibody responses against amyloid-beta without concomitant T cell responses. *Vaccine* 2006; 24: 6321–6331.
125. Jennings GT, Bachmann MF. The coming of age of virus-like particle vaccines. *Biol Chem* 2008; 389: 521–536.
126. Movsesyan N, Ghochikyan A, Mkrtichyan M, et al. Reducing AD-like pathology in 3xTg-AD mouse model by DNA epitope vaccine–a novel immunotherapeutic strategy. *PLoS One* 2008; 3: e2124.
127. Lemere CA, Spooner ET, Leverone JF, et al. Intranasal immunotherapy for the treatment of Alzheimer's disease: *Escherichia coli* LT and LT(R192G) as mucosal adjuvants. *Neurobiol Aging* 2002; 23: 991–1000.
128. Seabrook TJ, Thomas K, Jiang L, et al. Dendrimeric Abeta1-15 is an effective immunogen in wildtype and APP-tg mice. *Neurobiol Aging* 2007; 28: 813–823.
129. Nikolic WV, Bai Y, Obregon D, et al. Transcutaneous beta-amyloid immunization reduces cerebral beta-amyloid deposits without T cell infiltration and microhemorrhage. *Proc Natl Acad Sci U S A* 2007; 104: 2507–2512.
130. Frackowiak J, Wisniewski HM, Wegiel J, et al. Ultrastructure of the microglia that phagocytose amyloid and the microglia that produce beta-amyloid fibrils. *Acta Neuropathol* 1992; 84: 225–233.
131. Jucker M, Heppner FL. Cerebral and peripheral amyloid phagocytes–an old liaison with a new twist. *Neuron* 2008; 59: 8–10.
132. Butovsky O, Kunis G, Koronyo-Hamaoui M, Schwartz M. Selective ablation of bone marrow-derived dendritic cells increases amyloid plaques in a mouse Alzheimer's disease model. *Eur J Neurosci* 2007; 26: 413–416.
133. El KJ, Toft M, Hickman SE, et al. Ccr2 deficiency impairs microglial accumulation and accelerates progression of Alzheimer-like disease. *Nat Med* 2007; 13: 432–438.
134. Scholtzova H, Kascsak RJ, Bates KA, et al. Induction of Toll-like receptor 9 signaling as a method for ameliorating Alzheimer's disease related pathology. *J Neurosci* 2009; 29: 1846–1854.
135. Tahara K, Kim HD, Jin JJ, et al. Role of toll-like receptor signalling in Abeta uptake and clearance. *Brain* 2006; 129: 3006–3019.
136. Obregon D, Hou H, Bai Y, et al. CD40L disruption enhances Abeta vaccine-mediated reduction of cerebral amyloidosis while minimizing cerebral amyloid angiopathy and inflammation. *Neurobiol Dis* 2008; 29: 336–353.
137. Town T, Laouar Y, Pittenger C, et al. Blocking TGF-beta-Smad2/3 innate immune signaling mitigates Alzheimer-like pathology. *Nat Med* 2008; 14: 681–687.
138. Glabe CG. Structural classification of toxic amyloid oligomers. *J Biol Chem* 2008; 283: 29 639–29 643.
139. Klybin I, Betts V, Blennow K, et al. Aβ dimer-containing human cerebrospinal fluid disrupts synaptic plasticity: prevention by systemic passive immunization. *J Neurosci* 2008; 28: 4231–4237.
140. Lee EB, Leng LZ, Zhang B, et al. Targeting amyloid-beta peptide (Abeta) oligomers by passive immunization with a conformation-selective monoclonal antibody improves learning and memory in Abeta precursor protein (APP) transgenic mice. *J Biol Chem* 2006; 281: 4292–4299.
141. Moretto N, Bolchi A, Rivetti C, et al. Conformation-sensitive antibodies against Alzheimer amyloid-beta by immunization with a thioredoxin-constrained B-cell epitope peptide. *J Biol Chem* 2007; 282: 11 436–11 445.
142. Mamikonyan G, Necula M, Mkrtichyan M, et al. Anti-A beta 1–11 antibody binds to different beta-amyloid species, inhibits fibril formation, and disaggregates preformed fibrils but not the most toxic oligomers. *J Biol Chem* 2007; 282: 22 376–22 386.
143. Lambert MP, Velasco PT, Chang L, et al. Monoclonal antibodies that target pathological assemblies of Abeta. *J Neurochem* 2007; 100: 23–35.
144. Lambert MP, Velasco PT, Viola KL, Klein WL. Targeting generation of antibodies specific to conformational epitopes of amyloid beta-derived neurotoxins. *CNS Neurol Disord Drug Targets* 2009; 8: 65–81.
145. Puzzo D, Privitera L, Leznik E, et al. Picomolar amyloid-beta positively modulates synaptic plasticity and memory in hippocampus. *J Neurosci* 2008; 28: 14 537–14 545.

146. Giuffrida ML, Caraci F, Pignataro B, et al. Beta-amyloid monomers are neuroprotective. *J Neurosci* 2009; 29: 10 582–10 587.
147. Wisniewski T, Prelli F, Scholtzova H, et al. Immunotherapy targeting abnormal protein conformation. *Alzheimers Dement* 2009; 4: 113.
148. Wisniewski T, Sigurdsson EM. Therapeutic approaches for prion and Alzheimer's diseases. *FEBS J* 2007; 274: 3784–3798.
149. Goni F, Prelli F, Schreiber F, et al. High titers of mucosal and systemic anti-PrP antibodies abrogates oral prion infection in mucosal vaccinated mice. *Neuroscience* 2008; 153: 679–686.
150. Gotz J, Probst A, Spillantini MG, et al. Somatodendritic localization and hyperphosphorylation of tau protein in transgenic mice expressing the longest human brain tau isoform. *EMBO J* 1995; 14: 1304–1313.
151. Probst A, Gotz J, Wiederhold KH, et al. Axonopathy and amyotrophy in mice transgenic for human four-repeat tau protein. *Acta Neuropathol* 2000; 99: 469–481.
152. Spittaels K, Van den HC, Van Dorpe J, et al. Prominent axonopathy in the brain and spinal cord of transgenic mice overexpressing four-repeat human tau protein. *Am J Pathol* 1999; 155: 2153–2165.
153. Brion JP, Tremp G, Octave JN. Transgenic expression of the shortest human tau affects its compartmentalization and its phosphorylation as in the pretangle stage of Alzheimer's disease. *Am J Pathol* 1999; 154: 255–270.
154. Ishihara T, Hong M, Zhang B, et al. Age-dependent emergence and progression of a tauopathy in transgenic mice overexpressing the shortest human tau isoform. *Neuron* 1999; 24: 751–762.
155. Higuchi M, Ishihara T, Zhang B, et al. Transgenic mouse model of tauopathies with glial pathology and nervous system degeneration. *Neuron* 2002; 35: 433–446.
156. Boutajangout A, Leroy K, Touchet N, et al. Increased tau phosphorylation but absence of formation of neurofibrillary tangles in mice double transgenic for human tau and Alzheimer mutant (M146L) presenilin-1. *Neurosci Lett* 2002; 318: 29–33.
157. Boutajangout A, Authelet M, Blanchard V, et al. Characterisation of cytoskeletal abnormalities in mice transgenic for wild-type human tau and familial Alzheimer's disease mutants of APP and presenilin-1. *Neurobiol Dis* 2004; 15: 47–60.
158. Terwel D, Lasrado R, Snauwaert J, et al. Changed conformation of mutant Tau-P301L underlies the moribund tauopathy, absent in progressive, nonlethal axonopathy of Tau-4R/2N transgenic mice. *J Biol Chem* 2005; 280: 3963–3973.
159. Lewis J, McGowan E, Rockwood J, et al. Neurofibrillary tangles, amyotrophy and progressive motor disturbance in mice expressing mutant (P301L) tau protein. *Nat Genet* 2000; 25: 402–405.
160. Allen B, Ingram E, Takao M, et al. Abundant tau filaments and nonapoptotic neurodegeneration in transgenic mice expressing human P301S tau protein. *J Neurosci* 2002; 22: 9340–9351.
161. Gotz J, Tolnay M, Barmettler R, et al. Oligodendroglial tau filament formation in transgenic mice expressing G272V tau. *Eur J Neurosci* 2001; 13: 2131–2140.
162. Tatebayashi Y, Miyasaka T, Chui DH, et al. Tau filament formation and associative memory deficit in aged mice expressing mutant (R406W) human tau. *Proc Natl Acad Sci U S A* 2002; 99: 13 896–13 901.
163. Lim F, Hernandez F, Lucas JJ, et al. FTDP-17 mutations in tau transgenic mice provoke lysosomal abnormalities and tau filaments in forebrain. *Mol Cell Neurosci* 2001; 18: 702–714.
164. Schindowski K, Bretteville A, Leroy K, et al. Alzheimer's disease-like tau neuropathology leads to memory deficits and loss of functional synapses in a novel mutated tau transgenic mouse without any motor deficits. *Am J Pathol* 2006; 169: 599–616.

MOUNT SINAI JOURNAL OF MEDICINE 77:32–42, 2010

Neuropathology of Alzheimer's Disease

Daniel P. Perl, MD

Neuropathology Division, Mount Sinai School of Medicine, New York, NY

OUTLINE

GROSSLY VISIBLE LESIONS
MICROSCOPIC CHANGES
  Neurofibrillary Tangles
  Senile Plaques
  Vascular Amyloid Deposition
  Granulovacuolar Degeneration
  Eosinophilic Rodlike Inclusions
  Synaptic Loss
EMERGING PICTURE OF NEUROPATHOLOGY IN EARLY PHASES OF ALZHEIMER'S DISEASE
ALZHEIMER'S DISEASE AND ITS ASSOCIATED CONDITIONS
CONCLUSION

ABSTRACT

Alois Alzheimer first pointed out that the disease which would later bear his name has a distinct and recognizable neuropathological substrate. Since then, much has been added to our understanding of the pathological lesions associated with the condition. The 2 primary cardinal lesions associated with Alzheimer's disease are the neurofibrillary tangle and the senile plaque. The neurofibrillary tangle consists of abnormal accumulations of abnormally phosphorylated tau within the perikaryal cytoplasm of certain neurons. The senile plaque consists of a central core of beta-amyloid, a 4-kD peptide, surrounded by abnormally configured neuronal processes or neurites. Other neuropathological lesions are encountered in cases of Alzheimer's disease, but the disease is defined and recognized by these 2 cardinal lesions. Other lesions include poorly understood changes such as granulovacuolar degeneration and eosinophilic rodlike bodies (Hirano bodies). The loss of synaptic components is a change that clearly has a significant impact on cognitive function and represents another important morphological alteration. It is important to recognize that distinguishing between Alzheimer's disease, especially in its early stages, and normal aging may be very difficult, particularly if one is examining the brains of patients who died at an advanced old age. It is also noted that instances of pure forms of Alzheimer's disease, in the absence of other coexistent brain disease processes, such as infarctions or Parkinson's disease–related lesions, are relatively uncommon, and this must be taken into account by researchers who employ postmortem brain tissues for research. ***Mt Sinai J Med* 77:32–42, 2010.** © *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Daniel P. Perl
Neuropathology Division
Mount Sinai School of Medicine
New York, NY
Email: daniel.perl@mssm.edu

Key Words: Alzheimer's disease, beta-amyloid, neurofibrillary tangle, neuropathology, senile plaque.

Since the original description by Alois Alzheimer of the disease that bears his name, the autopsy continues to represent the means by which the definitive diagnosis is made. In the absence of a valid biological marker, this situation continues. Furthermore, in the absence of a truly complete animal model, Alzheimer's disease remains a disorder of humans, and studies of actual patients with the condition and their brain tissues represent a critical route by which the condition can truly be characterized and understood. It is not likely that this situation will change in the near future, and thus the work of numerous neuropathologists and other neuroscientists who study human autopsy–derived specimens will remain vital to further scientific progress on this most important age-related condition. Here we describe the neuropathological features of the disease and some of the associated issues that arise for individuals who plan research on the disorder.

GROSSLY VISIBLE LESIONS

At the autopsy table, the brain of a patient with Alzheimer's disease does not show any grossly

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20157

© 2010 Mount Sinai School of Medicine

apparent alteration that can be considered to be diagnostic. By and large, any changes that are encountered are considered to be entirely nonspecific in nature and show extensive overlap with findings that are also noted in brain specimens derived from elderly individuals who had displayed normal cognitive function during life. Brain specimens obtained from most cases of Alzheimer's disease show a modest degree of cerebral cortical atrophy primarily but not exclusively involving the frontotemporal association cortex. In particular, this atrophy tends to spare primary motor, sensory, and visual areas. However, among elderly subjects, especially those of advanced age, there is considerable overlap between brain weight as well as measures of cerebral cortical thickness between age-matched individuals with normal cognitive function and those with Alzheimer's disease.[1] Accordingly, it becomes difficult, if not impossible, to take the brain weight or cerebral cortical thickness of any individual specimen derived from an elderly individual and, by that value alone, determine whether the person had Alzheimer's disease or not. However, when one is examining cases of Alzheimer's disease that are of presenile onset (arbitrarily set by most as an onset before 65 years of age), comparisons of the brain weight or the degree of cerebral cortical atrophy with those of age-matched controls usually reveal a clear and obvious difference upon gross inspection of the brain.

At the autopsy table, the brain of a patient with Alzheimer's disease does not show any grossly apparent alteration that can be considered to be diagnostic.

In cases of Alzheimer's disease, there is an associated loss of brain tissue that generally leads to a symmetrical dilation of the lateral ventricles (hydrocephalus ex vacuo). Again, by examining the lateral ventricles of advanced elderly individuals, one is not able to distinguish with certainty age-matched normals from cases with Alzheimer's disease. However, significant atrophy of the hippocampus, with an associated selective dilatation of the adjacent temporal horn of the lateral ventricle, represents a reasonably reliable clue upon dissection of the brain, suggesting to the neuropathologist that the case will ultimately show microscopic evidence of Alzheimer's disease. In a similar fashion, the absence of this finding is strongly suggestive that other explanations for the underlying cause of dementia will need to be sought.

MICROSCOPIC CHANGES

It is upon the histological examination of the brain specimen that a definitive diagnosis of Alzheimer's disease can be made. This diagnosis is based on the identification of a series of morphological abnormalities that are distributed in a rather stereotyped pattern. It is important to realize that virtually any of the alterations upon which the neuropathological diagnosis of Alzheimer's disease is made may also be seen, to some degree, in the brains of elderly individuals who during life had shown normal cognitive function. Often it requires the recognition of the extent of involvement in certain areas of the brain by such lesions for the neuropathologist to be able to declare that Alzheimer's disease was the cause of a patient's cognitive impairment. At times, this may involve subtle distinctions that require considerable experience and expertise, and at other times, the honest neuropathologist may be required to state that this diagnosis can or cannot be rendered with only a degree of probability and not certainty. This is particularly true when one is examining the brains of subjects dying within the oldest-old age range. In individuals dying at an extremely old age, the overlap of neuropathological changes between those found with severe dementia and those with intact cognitive function is extensive, and such distinctions may be very difficult. Indeed, correlations between the extent and distribution of such lesions are currently undergoing detailed scientific study using brain specimens derived from individuals in large cohorts of advanced elderly subjects who underwent rigorous longitudinal neuropsychological studies.[2]

It is upon the histological examination of the brain specimen that a definitive diagnosis of Alzheimer's disease can be made.

Neurofibrillary Tangles

In his pioneering article,[3] Alois Alzheimer, describing a single case of the disease that would eventually bear his name, noted the presence of abnormal fibrous inclusions within the perikaryal cytoplasm of pyramidal neurons. These inclusions are called Alzheimer neurofibrillary tangles and to this day are considered a cardinal microscopic lesion associated with the disease and a requirement for making the pathological diagnosis. Neurofibrillary tangles are very difficult to see with the traditional morphological stain used by pathologists, hematoxylin and eosin. By and large, one of a variety of silver impregnation staining techniques, such as the modified Bielschowski or Gallyas technique, or the fluorochrome dye thioflavin S is typically employed to visualize neurofibrillary tangles (Figures 1 and 2). Additionally, there are a number of immunohistochemical approaches used to visualize neurofibrillary tangles. These have mostly employed antibodies directed against abnormally phosphorylated *tau* (discussed later)[4–6] (Figure 3). Because these techniques require either specialized equipment (the thioflavin S stain requires the use of a fluorescence microscope with specialized excitatory and barrier filters) or experienced histotechnologists (in the case of silver impregnation stains), most anatomic pathologists in general practice lack the capability to properly evaluate brain specimens submitted for the diagnosis of Alzheimer's disease. Such specimens are best referred to specialized neuropathology laboratories where the necessary experience and facilities are in place.

> *Alois Alzheimer, describing the first patient with the disease that would eventually bear his name, noted the presence of abnormal fibrous inclusions within the perikaryal cytoplasm of pyramidal neurons. These Alzheimer neurofibrillary tangles to this day are considered a cardinal microscopic lesion associated with the disease.*

With these special stains, within neurons with a pyramidal shape to the perikaryon, such as those of the cornu ammonis 1 (CA1) sector of the hippocampus and the layer V neurons in areas of the association cortex, the neurofibrillary tangles appear as parallel, thickened fibrils that surround the nucleus and extend toward the apical dendrite (Figure 4). When the neurofibrillary tangle occurs within a neuron with a more rounded configuration (eg, neurons within the substantia nigra and locus ceruleus), the inclusion appears as interweaving swirls of fibers, and here it is called a globoid neurofibrillary tangle.

The nature of neurofibrillary tangles has been extensively studied over the past several decades, and much has been learned about its structural components. Ultrastructurally, the neurofibrillary tangle is composed of abnormal fibrils measuring 10 nm in diameter that occur in pairs and are wound in a helical fashion with a regular periodicity of 80 nm.[7,8] On the basis of these observations, such structures are generally called paired helical filaments. The primary constituent of the neurofibrillary tangle is the microtubule-associated protein *tau*. The *tau* within neurofibrillary tangles is abnormally phosphorylated with phosphate groups attached to very specific sites

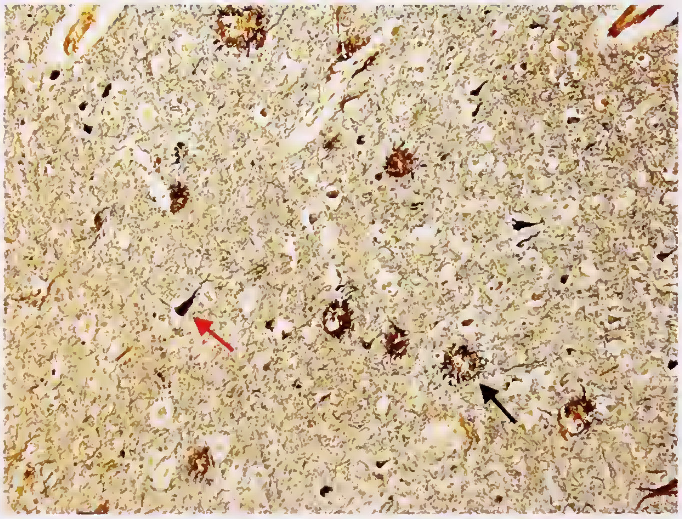


Fig 1. Photomicrograph of the temporal cortex of a patient with Alzheimer's disease (modified Bielschowski stain; original magnification, 40×). Numerous senile (neuritic) plaques (black arrow) and neurofibrillary tangles (red arrow) are shown.

DOI:10.1002/MSJ

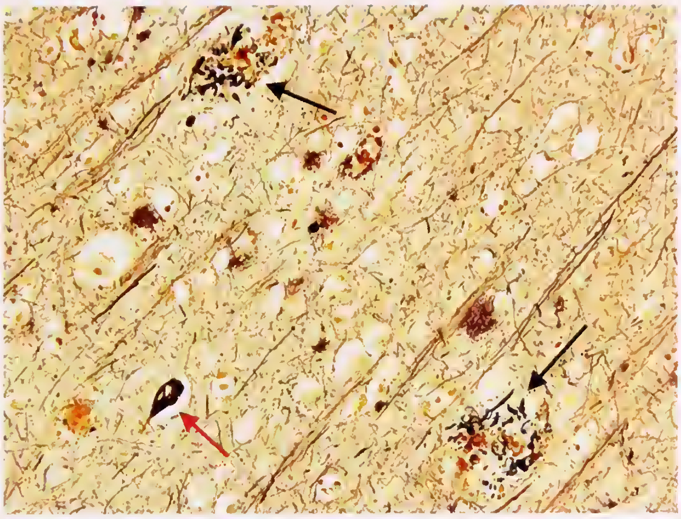


Fig 2. Photomicrograph of the temporal cortex of a patient with Alzheimer's disease (modified Bielschowski stain; original magnification, 100×). Numerous senile (neuritic) plaques (black arrows) and neurofibrillary tangles (red arrow) are shown.

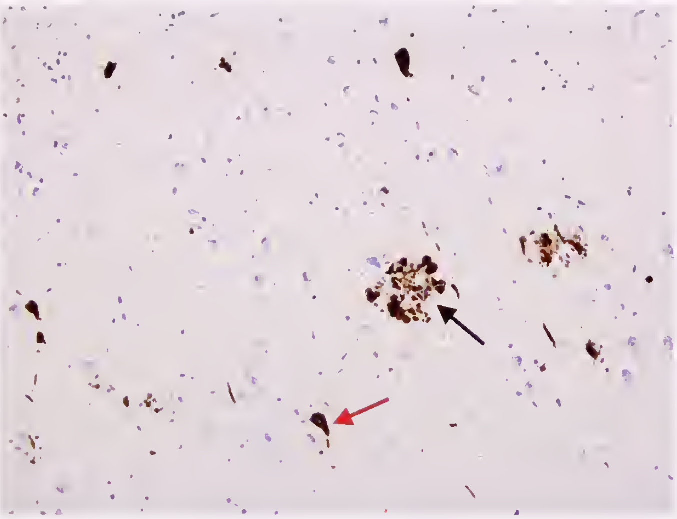


Fig 3. Temporal cortex of a patient with Alzheimer's disease (immunohistochemical stain; original magnification, 100×): the microscopic appearance of an immunohistochemical preparation using an antibody directed against abnormally phosphorylated tau (TG-3; a gift from Dr. P. Davies). This antibody prominently decorates neurofibrillary tangles (red arrow) and swollen dystrophic neurites (neuronal processes) that form the outer rim of the senile (neuritic) plaques (black arrow).

on the molecule.[9] There are a number of other protein constituents associated with the neurofibrillary tangle, such as ubiquitin,[10,11] cholinesterases,[12] and beta-amyloid 4 (βA4; discussed later),[13] but *tau* is considered to be the critical constituent of most of these structures.

The pattern of distribution of neurofibrillary tangles present in cases of Alzheimer's disease is, for the most part, rather stereotyped and predictable. Severe involvement is seen in the layer II neurons of the entorhinal cortex, the CA1 and subicular regions of the hippocampus, the amygdala, and the deeper layers (layers III, V, and superficial VI) of the neocortex.[14] Studies have shown that the extent and distribution of neurofibrillary tangles in cases of Alzheimer's disease correlate with both the degree of dementia and the duration of illness,[15,16] and this suggests that these abnormalities

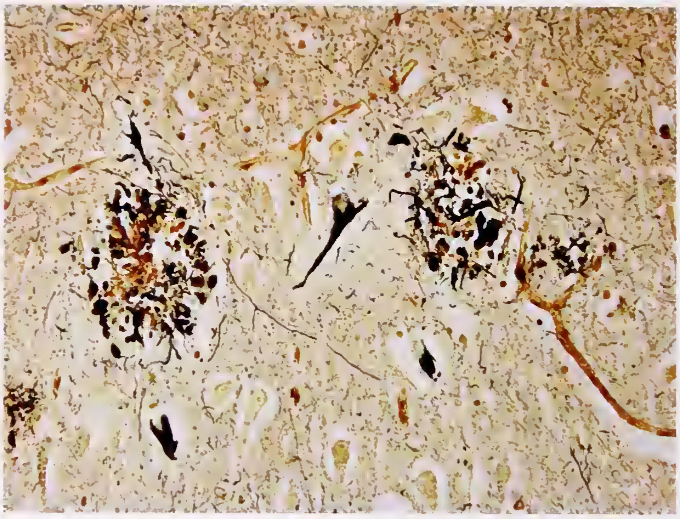


Fig 4. Photomicrograph of the temporal cortex of a patient with Alzheimer's disease (modified Bielschowski stain; original magnification, 400×). Two senile (neuritic) plaques with a neurofibrillary tangle between them are shown.

do have some direct impact on the functioning capacity of the brain. However, it is also clear that other factors contribute to the production of the clinical features of the disease. Although the neurofibrillary tangle is considered a cardinal histopathological feature of Alzheimer's disease, it should be kept in mind that this neuropathological lesion may also be encountered in association with many other disease states.[17] These include disorders such as postencephalitic parkinsonism, posttraumatic dementia or dementia pugalistica, type C Niemann-Pick disease, and amyotrophic lateral sclerosis/parkinsonism dementia complex of Guam. It remains unclear why diseases with such a wide range of underlying etiological mechanisms should all show this one particular neuropathological abnormality.

Studies have shown that the extent and distribution of neurofibrillary tangles in cases of Alzheimer's disease correlate with both the degree of dementia and the duration of illness.

Senile Plaques

The other cardinal pathological lesion encountered in patients suffering from Alzheimer's disease is the senile or neuritic plaque. Senile plaques are complex structures that are defined by the presence of a central core accumulation of a 4-kD protein with a beta-pleated sheet configuration called βA4.[18–20] The predominant beta-pleated sheet configuration of this protein confers its ability to bind the planar dye Congo red and produce birefringence when illuminated by polarized light; it thus conforms to the physical definition of an amyloid.

The other cardinal pathological lesion encountered in patients suffering from Alzheimer's disease is the senile or neuritic plaque.

The brains of aged individuals and cases of Alzheimer's disease may also contain several forms of βA4-containing plaques, and at times, the nomenclature employed for these lesions in the literature can be confusing. The senile or neuritic plaque has a central core of βA4 protein arranged in a radial fashion and is surrounded by a corona of abnormally formed neurites (or neuronal processes, either dendrites or axons). These abnormal or dystrophic neurites stain strongly with the same silver impregnation stains used to identify the neurofibrillary tangles (Figures 2 and 5), and ultrastructurally, these structures contain dense bodies, membranous profiles, and packets of paired helical filaments. In the periphery of the neuritic plaque, one commonly encounters 1 to several microglial cells and, less frequently, reactive astrocytes. Whether these microglial cells are actively involved in a neuroinflammatory pathogenetic cascade or are reacting to the presence

DOI:10.1002/MSJ

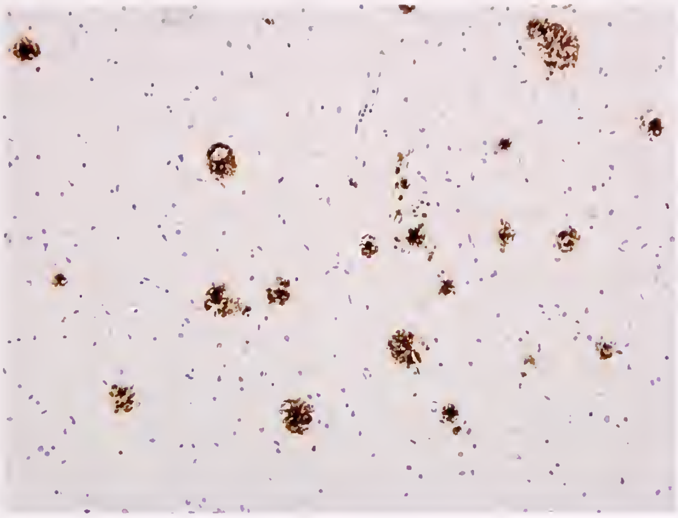


Fig 5. Temporal cortex of a patient with Alzheimer's disease (immunohistochemical stain; original magnification, 100×): the microscopic appearance of an immunohistochemical preparation using an antibody directed against the components of beta-amyloid (4G8; a gift from Dr. Robakis). This antibody selectively decorates the numerous senile plaques present in this case of advanced Alzheimer's disease and demonstrates the extent of amyloid accumulation that one may encounter in the terminal phases of the disease.

of constituents within these lesions remains a matter of intense debate.

Through the use of immunohistochemical techniques with antibodies raised against portions of the βA4 molecule, it has been recognized that focal diffuse deposits of this amyloid protein may occur in the cerebral cortex in the absence of accompanying dystrophic neurites.[21] Such deposits of βA4 will also stain with the aforementioned silver-based stains and especially with enhanced silver stains such as those of the Hedreen, Tago, and Campbell-Switzer methods, and they are called diffuse plaques. Diffuse plaques are very commonly encountered in the brains of elderly individuals and can be seen in relatively large numbers in the absence of any associated evidence of cognitive impairment.[22–24]

A third form of plaque, consisting of a dense core of βA4 that does not show any accompanying dystrophic neurites, has also been identified. Such plaques have been called burned-out plaques and end-stage plaques and are considered to be the remnants of what were once neuritic plaques.[25] Because neuropathological observations of such lesions are, by their very nature, cross-sectional, these interpretations should be best thought of as speculation.

The βA4 protein is derived from a larger amyloid precursor protein, a highly conserved transmembrane glycoprotein.[20] The βA4 portion of the precursor protein is formed through the action of 2 secretases, which split the amino and carboxyl terminals of the 4-kD segment, which is then released and accumulates within brain tissues.[26–28] The carboxyl end cleavage is ragged, with longer forms (having a total of 42 or 43 amino acids) tending to be deposited within senile plaques and a shorter form (containing 40 amino acids) tending to accumulate within the leptomeningeal and cerebral cortical and cerebellar blood vessels (congophilic angiopathy; discussed later).[29] The central cores of senile plaques have been shown to contain several other proteins, such as heparan sulfate glycoproteins,[30–32] apolipoprotein E,[33–35] complement proteins,[36–38] and alpha-1-antichymotrypsin.[39]

Vascular Amyloid Deposition (Congophilic Angiopathy)

In addition to βA4 accumulating in the central core of neuritic plaques, this protein also tends to deposit in the walls of the cerebral cortical blood vessels. The protein accumulates in the walls of small arteries and arterioles of the leptomeninges and within the gray matter of the cerebral cortex. This lesion is often called congophilic angiopathy, which reflects the ability of the Congo red dye to attach to the lesion. These accumulations of βA4 within the vessel walls do not appear to clog the vascular lumina or otherwise interfere with the function of these vessels.

However, when the degree of vascular involvement is severe, there is a tendency for spontaneous vascular rupture leading to a focal accumulation of blood in the brain tissues. Such hemorrhages are generally not encountered in and around the lenticular nuclei and thalamus, such as is seen as a result of uncontrolled systemic hypertension. These hemorrhages tend to occur in the white matter of the frontal and/or occipital poles, are often small and multiple, and may be microscopic in size. When they are large (this is a relatively rare situation), they may be multiple and are commonly called lobar hemorrhages. Although rare, such lobar hemorrhages represent one of the few fatal intracerebral complications of Alzheimer's disease.

Granulovacuolar Degeneration

Simchowitz[40] first identified granulovacuolar degeneration in 1911. This is a poorly understood lesion that consists of an intraneuronal cluster of small vacuoles measuring 2 to 4 µm in diameter, each containing a small, dense basophilic granule that typically measures approximately 1 µm in diameter. These granule-containing vacuoles are encountered almost exclusively within the perikaryal cytoplasm of pyramidal neurons of the hippocampus, typically in the neurons that form the junction between the cornu ammonis 2 (CA2) and CA1 sectors of Ammon's horn. They are also seen in lesser numbers in the neurons of the remainder of the CA1 hippocampal region. Ultrastructurally, they are described as membrane-bound structures with a clear vacuolar matrix containing an electron-dense central granule.[41] The central granules stain intensely with silver impregnation stains and with antibodies directed against phosphorylated neurofilaments, tubulin, *tau*, and ubiquitin.[42–44] Little is known about the nature of these lesions or their significance. They are seen in brain specimens derived from elderly individuals with normal cognitive function, but 2 studies have shown that large numbers of such lesions in the boundary zone between the CA1 and CA2 regions of the caudal aspect of the hippocampus correlate well with a diagnosis of Alzheimer's disease.[45,46]

Eosinophilic Rodlike Inclusions (Hirano Bodies)

Eosinophilic rodlike inclusions, or Hirano bodies, are intensely eosinophilic perineuronal lesions encountered within the CA1 region of the hippocampus. These lesions were first identified by Asao Hirano in his studies of amyotrophic lateral sclerosis/parkinsonism-dementia complex of Guam, and they are now commonly called Hirano bodies.[47] This disease is mostly confined to the indigenous native population living on Guam, and in this setting, large numbers of Hirano bodies are commonly encountered. Identical inclusions in much smaller numbers are also noted in some, but not all, cases of Alzheimer's disease. Hirano bodies may also occasionally be encountered in the brains of normal aged individuals with intact cognition. Hirano bodies have a very characteristic ultrastructural appearance consisting of parallel fibers that interweave in a very regular crossing pattern reminiscent of the appearance of a tweed fabric. Immunohistochemical studies indicate the presence of actin, tropomyosin, and vinculin within these bodies.[48–50] The Hirano body is currently considered to be a nonspecific lesion of unknown significance.

Synaptic Loss

Masliah and coworkers[51–54] as well as Scheff and coworkers[55–59] have demonstrated that a substantial loss in synaptic profiles occurs in the brains of patients with Alzheimer's disease. This has been investigated by the quantification of immunohistochemical markers directed against synaptic proteins, such as synaptophysin, and by quantitative electron microscopy. In comparison with normal controls, Masliah and coworkers have shown a 45% loss of the extent of staining of presynaptic boutons in cases of Alzheimer's disease. These workers have argued that this loss of the critical element for neuron-to-neuron communication constitutes the major morphological counterpart to cognitive loss in Alzheimer's disease. Their data indicate that this loss correlates strongly with the degree of functional impairment.[52]

Masliah and coworkers have shown a 45% loss of the extent of staining of presynaptic boutons in cases of Alzheimer's disease in comparison with normal controls and have argued that this loss of the critical element for neuron-to-neuron communication constitutes the major morphological counterpart to cognitive loss in Alzheimer's disease.

EMERGING PICTURE OF NEUROPATHOLOGY IN EARLY PHASES OF ALZHEIMER'S DISEASE

Over the years, most of the neuropathological literature related to descriptions of Alzheimer's disease and its clinicopathological correlation has tended to involve detailed examinations of the end-stage or terminal phases of the disease. Because patients with Alzheimer's disease do not typically die as a direct result of the disease itself, many patients with the disorder, especially if they had a relatively young onset of disease, will tend to survive into its more advanced stages. Such patients typically tend to survive for many years in a severely impaired state until a superimposed pneumonia or other infection emerges to ultimately serve as the cause of death. Accordingly, the brain specimen obtained from such an end-stage case will demonstrate an extensive burden of neurodegenerative lesions in a widespread distribution. Specimens derived from such terminal cases will typically show a very heavy burden of both neurofibrillary tangles and senile plaques. In such cases, it is virtually impossible to dissect out meaningful clinicopathological correlational relationships.

It is clear that Alzheimer's disease is a slowly progressive disorder whose lesions accumulate within the brain over a period of many years. How long it takes before a sufficient extent of neuropathological damage occurs to produce a degree of functional impairment that might allow a clinical diagnosis to be considered is not known, but it is likely to take many years, if not decades, for this to occur. Additional years are likely needed for the further progression from the early stages to the middle stages of the disorder. Presumably, the progressive accumulation of a burden of the aforementioned neuropathological lesions underlies this clinical progression. Recognizing that advanced age, with its inherent frailties and multiple medical comorbidities, leads to a high death rate, we anticipate that many individuals of advanced age will die of unrelated causes while they are also passing through the early clinical and even preclinical stages of Alzheimer's disease. It can be argued that the neuropathological changes found in such preclinical cases of Alzheimer's disease might well have been considered to be within the range of normal aging.

This has led some to begin thinking about the progressive stages of involvement of the brain by Alzheimer's disease and to speculate on what the earliest phases of the disease might look like. In 1991, neuroanatomists Eva and Heiko Braak published a proposed sequence of progression of the neuropathology of Alzheimer's disease, breaking the disorder down into 6 stages with increasing involvement of the brain (the so-called Braak and Braak stages). Braak and Braak stages 1 and 2 show selective involvement by neurofibrillary tangles in the transentorhinal cortex. This is followed by stages 3 and 4 with increasing limbic lobe involvement, with the final 2 stages (stages 5 and 6) showing the more typical widespread pattern of involvement in the neocortex. These Braak and Braak stages were initially developed by the evaluation of the pattern of neurodegenerative changes (predominantly neurofibrillary tangles) present in a series of 83 brain specimens derived from elderly individuals. They devised the stages by mapping out the extent and distribution of lesions in these brain specimens, but no associated clinical data were available from the patients from whom the samples were obtained. Although it is clear that not all Alzheimer's disease patients progress precisely along the stages they described, the Braak and Braak stages do represent a useful concept and have provided a format for neuropathologists to use in evaluating the relative stage of development of the disease.

ALZHEIMER'S DISEASE AND ITS ASSOCIATED CONDITIONS

It is very common for both clinicians and researchers involved in studying Alzheimer's disease to approach a patient or brain specimen, for that matter, and assume that the diagnosis is solely Alzheimer's disease and nothing else. Although this may at times be true, more often the reality is that one or another pathological process is also present in the brain in addition to Alzheimer's disease. Indeed, certainly at autopsy, it is relatively uncommon to encounter Alzheimer's disease in its pure form, that is, in the absence of 1 or more coexisting conditions. This is particularly true when one is examining the brains of advanced elderly patients who suffered from dementia. It is very important to recognize this situation when one is designing experiments involving the use of postmortem brain tissues.

The most common coexisting condition with Alzheimer's disease is stroke and more specifically ischemic infarction, another very common

The most common coexisting condition with Alzheimer's disease is stroke and more specifically ischemic infarction.

DOI:10.1002/MSJ

age-related brain disease. Over the age of 75 years, ischemic infarctions of the brain are rather frequently encountered, and the prevalence of such lesions rises dramatically as the age at death increases. Accordingly, it is very common to encounter 1 or more such ischemic lesions in the brains of victims of Alzheimer's disease. Very often, clinical evidence of the onset of stroke that correlates with the presence of such lesions is missing. This can arise from a lack of a referral for a neurological evaluation or from the inability of the cognitively impaired individual to give a detailed history. Inevitably, the pathologist is left with a specimen demonstrating the full neuropathological picture of Alzheimer's disease, as described previously, plus superimposed ischemic infarct(s) typically involving the territory served by the middle cerebral arteries. Over the past 20+ years of examining the brains of deceased elderly patients who suffered from dementia within the Mount Sinai Brain Bank repository, approximately 20% of cases have shown evidence of both Alzheimer's disease and superimposed brain infarction. Such cases are generally referred to as mixed dementia, and one is often left unable to determine the relative contribution of the stroke(s) in producing the cognitive impairment, as opposed to the underlying Alzheimer's disease.

Another issue of comorbid pathological features is the common occurrence of neuropathological lesions of Alzheimer's disease coexisting with those of Parkinson's disease. Neuropathological studies of large numbers of patients who, during life, had a clinical diagnosis of Alzheimer's disease demonstrate not only the neuropathological features of this disease but also evidence of neuronal degeneration of the substantia nigra pars compacta with Lewy bodies in many remaining neurons. The later findings suggest a coexistent diagnosis of Parkinson's disease. In cases of this type, it is common to also see evidence of widely distributed cortical Lewy bodies, especially if appropriate immunohistochemical preparations are made. Such cases are now commonly diagnosed as suffering from dementia with Lewy bodies, and the criteria for both clinical and neuropathological diagnoses of this condition have been written and subsequently revised by consensus committees.[60,61] The boundary line between Alzheimer's disease and Parkinson's disease and the nosologic status of dementia with Lewy bodies remains a subject of debate, but this is an important instance in which the concept of the common age-related neurodegenerative diseases remaining in a pure state is challenged.

> *Neuropathological studies of large numbers of patients who, during life, had a clinical diagnosis of Alzheimer's disease also demonstrate [the characteristic lesions of Parkinson's disease, that is,] neuronal degeneration of the substantia nigra pars compacta with Lewy bodies in many remaining neurons.*

CONCLUSION

It is over 100 years since Alois Alzheimer first described the neurofibrillary tangles associated with Alzheimer's disease and introduced the notion that this condition has an underlying neuropathological substrate. Many more details have been added over the past century. Despite this progress, there are still important gaps in our understanding of the condition and the nature of the pathological processes taking place in the brain. Although the neuropathological diagnosis can be established with relative precision among those patients with a relatively young onset, such cases are relatively rare in comparison with the vast burden of disease encountered in individuals of advanced age. Here, the distinction from normal aging and the correlations between structure and function are very difficult to delineate and will require considerable additional study. The continued morphological study of human brain specimens derived from longitudinally assessed patients will remain the cornerstone for such research.

ACKNOWLEDGMENT

This work was supported by grants P50 AG-05 138 and P01 AG-02 219 from the National Institutes of Health.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Terry RD. Interrelations among the lesions of normal and abnormal aging of the brain. *Prog Brain Res* 1986; 70: 41–48.

2. Haroutunian V, Schnaider-Beeri M, Schmeidler J, et al. Role of the neuropathology of Alzheimer disease in dementia in the oldest-old. *Arch Neurol* 2008; 65: 1211–1217.
3. Alzheimer A. Ueber eine eigenartige erkrankung der hirnrinde. *Z Gesamte Neurol Psychiatr* 1907; 18: 177–179.
4. Dickson DW, Farlo J, Davies P, et al. Alzheimer's disease. A double-labeling immunohistochemical study of senile plaques. *Am J Pathol* 1988; 132: 86–101.
5. Yen SH, Dickson DW, Crowe A, et al. Alzheimer's neurofibrillary tangles contain unique epitopes and epitopes in common with the heat-stable microtubule associated proteins tau and MAP2. *Am J Pathol* 1987; 126: 81–91.
6. Iwatsubo T, Hasegawa M, Ihara Y. Neuronal and glial tau-positive inclusions in diverse neurologic diseases share common phosphorylation characteristics. *Acta Neuropathol (Berl)* 1994; 88: 129–136.
7. Kidd M. Paired helical filaments in electron microscopy of Alzheimer's disease. *Nature* 1963; 197: 192–193.
8. Wisniewski HM, Narang HK, Terry RD. Neurofibrillary tangles of paired helical filaments. *J Neurol Sci* 1976; 27: 173–181.
9. Lee VMY, Balin BJ, Otvos L Jr, Trojanowski JQ. A major subunit of paired helical filaments and derivatized forms of normal tau. *Science* 1991; 251: 675–678.
10. Perry G, Friedman R, Shaw G, Chau C. Ubiquitin is detected in neurofibrillary tangles and senile plaque neurites of Alzheimer's disease brains. *Proc Natl Acad Sci U S A* 1987; 84: 3033–3036.
11. Love S, Saitoh T, Quijada S, et al. Alz-50, ubiquitin and tau immunoreactivity of neurofibrillary tangles, Pick bodies and Lewy bodies. *J Neuropathol Exp Neurol* 1988; 47: 393–405.
12. Mesulam MM, Moran MA. Cholinesterases within neurofibrillary tangles related to age and Alzheimer's disease. *Ann Neurol* 1987; 22: 223–228.
13. Hyman BT, Van Hoesen GW, Beyreuther K, Masters CL. A4 amyloid protein immunoreactivity is present in Alzheimer's disease neurofibrillary tangles. *Neurosci Lett* 1989; 101: 352–355.
14. Morrison JH, Hof PR. Life and death of neurons in the aging brain. *Science* 1997; 278: 412–419.
15. Hof PR, Bierer LM, Purohit DP, et al. Neurofibrillary tangles correlate with dementia severity in Alzheimer's disease. *Arch Neurol* 1995; 52: 81–88.
16. Arriagada PV, Growdon JH, Hedley-White T, Hyman BT. Neurofibrillary tangles and not senile plaques parallel duration and severity of Alzheimer's disease. *Neurology* 1992; 42: 1681–1688.
17. Wisniewski K, Jervis GA, Moretz RC, Wisniewski HM. Alzheimer neurofibrillary tangles in diseases other than senile and presenile dementia. *Ann Neurol* 1979; 5: 288–294.
18. Masters CL, Multhaup G, Simms G, et al. Neuronal origin of a cerebral amyloid: neurofibrillary tangles of Alzheimer's disease contain the same protein as the amyloid of plaque cores and blood vessels. *EMBO J* 1985; 4: 2757–2763.
19. Beyreuther K, Masters CL. Nomenclature of amyloid A4 proteins and their precursors in Alzheimer's disease and Down's syndrome. *Neurobiol Aging* 1990; 11: 66–68.
20. Kang J, Lemaire HG, Unterbeck A, et al. The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature* 1987; 325: 733–736.
21. Yamaguchi H, Hirai S, Morimatsu M, et al. Diffuse type of senile plaques in the cerebellum of Alzheimer-type dementia demonstrated by beta protein immunostain. *Acta Neuropathol* 1989; 77: 314–319.
22. Gentleman SM, Bruton C, Allsop D, et al. A demonstration of the advantages of immunostaining in the quantification of amyloid plaque deposits. *Histochemistry* 1989; 92: 355–358.
23. Wolf DS, Gearing M, Snowdon DA, et al. Progression of regional neuropathology in Alzheimer disease and normal elderly: findings from the Nun study. *Alzheimer Dis Assoc Disord* 1999; 13: 226–231.
24. Morris JC, Storandt M, McKeel DW Jr, et al. Cerebral amyloid deposition and diffuse plaques in "normal" aging: evidence for presymptomatic and very mild Alzheimer's disease. *Neurology* 1996; 46: 707–719.
25. Wisniewski HM, Vorbrodt AW, Moretz RC, et al. Pathogenesis of neuritic (senile) and amyloid plaque formation. *Exp Brain Res* 1982; S5: 3–9.
26. Selkoe DJ, Yamazaki T, Citron M, et al. The role of APP processing and trafficking pathways in the formation of amyloid beta-protein. *Ann N Y Acad Sci* 1996; 777: 57–64.
27. Selkoe DJ. The cell biology of beta-amyloid precursor protein and presenilin in Alzheimer's disease. *Trends Cell Biol* 1998; 8: 447–453.
28. Selkoe DJ. Alzheimer's disease: genes, proteins, and therapy. *Physiol Rev* 2001; 81: 741–766.
29. Prelli F, Castano E, Glenner GG, Frangione B. Differences between vascular and plaque core amyloid in Alzheimer's disease. *J Neurochem* 1988; 51: 648–651.
30. Castillo GM, Ngo C, Cummings J, et al. Perlecan binds to the beta-amyloid proteins (A beta) of Alzheimer's disease, accelerates A beta fibril formation, and maintains A beta fibril stability. *J Neurochem* 1997; 69: 2452–2465.
31. Maresh GA, Erezyilmaz D, Murry CE, et al. Detection and quantitation of perlecan mRNA levels in Alzheimer's disease and normal aged hippocampus by competitive reverse transcription-polymerase chain reaction. *J Neurochem* 1996; 67: 1132–1144.
32. Snow AD, Nochlin D, Sekiguichi R, Carlson SS. Identification in immunolocalization of a new class of proteoglycan (keratan sulfate) to the neuritic plaques of Alzheimer's disease. *Exp Neurol* 1996; 138: 305–317.
33. Namba Y, Tomonaga M, Kawasaki H, et al. Apolipoprotein E immunoreactivity in cerebral amyloid deposits and neurofibrillary tangles in Alzheimer's disease and kuru plaque amyloid in Creutzfeldt-Jakob disease. *Brain Res* 1991; 541: 163–166.
34. Dickson TC, Saunders HL, Vickers JC. Relationship between apolipoprotein E and the amyloid deposits and dystrophic neurites of Alzheimer's disease. *Neuropathol Appl Neurobiol* 1997; 23: 483–491.
35. Nishiyama E, Iwamoto N, Ohwada J, Arai H. Distribution of apolipoprotein E in senile plaques in brains with Alzheimer's disease: investigation with the confocal laser scan microscope. *Brain Res* 1997; 750: 20–24.
36. Dickson DW, Rogers J. Neuroimmunology of Alzheimer's disease: a conference report. *Neurobiol Aging* 1992; 13: 793–798.

37. Rogers J, Cooper NR, Webster S, et al. Complement activation by beta-amyloid in Alzheimer disease. *Proc Natl Acad Sci U S A* 1992; 89: 10016–10020.
38. McGeer PL, Rogers J, McGeer EG. Neuroimmune mechanisms in Alzheimer disease pathogenesis. *Alzheimer Dis Assoc Disord* 1994; 8: 149–158.
39. Abraham CR, Shirahama T, Potter H. Alpha 1-antichymotrypsin is associated solely with amyloid deposits containing the beta-protein. Amyloid and cell localization of alpha 1-antichymotrypsin. *Neurobiol Aging* 1990; 11: 123–129.
40. Simchowicz T. Histologische studien über die senile dementz in histol. und histopathol. *Arbeiten Aber Grosshirnrinde* 1911; 4: 267–260.
41. Hirano A, Dembitzer HM, Kurland LT, Zimmerman HM. The fine structure of some intraganglionic alterations. Neurofibrillary tangles, granulovacuolar bodies and "rod-like" structures as seen in Guam amyotrophic lateral sclerosis and parkinsonism-dementia complex. *J Neuropathol Exp Neurol* 1968; 27: 167–182.
42. Kahn J, Anderton BH, Probst A, et al. Immunohistological study of granulovacuolar degeneration using monoclonal antibodies to neurofilaments. *J Neurol Neurosurg Psychiatry* 1985; 48: 924–926.
43. Dickson DW, Ksiezak-Reding H, Davies P, Yen SH. A monoclonal antibody that recognizes a phosphorylated epitope in Alzheimer neurofibrillary tangles, neurofilaments and tau proteins immunostains granulovacuolar degeneration. *Acta Neuropathol (Berl)* 1987; 73: 254–258.
44. Ikegami K, Kimura T, Katsuragi S, et al. Immunohistochemical examination of phosphorylated tau in granulovacuolar degeneration granules. *Psychiatry Clin Neurosci* 1996; 50: 137–140.
45. Tomlinson BE, Kitchener D. Granulovacuolar degeneration of hippocampal pyramidal cells. *J Pathol* 1972; 106: 165–185.
46. Ball MJ, Lo L. Granulovacuolar degeneration in the ageing brain and in dementia. *J Neuropathol Exp Neurol* 1977; 36: 474–487.
47. Hirano A. Hirano bodies and related neuronal inclusions. *Neuropathol Appl Neurobiol* 1994; 20: 3–11.
48. Galloway PG, Perry G, Gambetti P. Hirano bodies contain actin and actin-associated proteins. *J Neuropathol Exp Neurol* 1987; 46: 185–199.
49. Galloway PG, Perry G, Kosik K, Gambetti P. Hirano bodies contain tau protein. *Brain Res* 1987; 403: 337–340.
50. Goldman JE. The association of actin with Hirano bodies. *J Neuropathol Exp Neurol* 1983; 42: 146–152.
51. Masliah E, Terry R. The role of synaptic proteins in the pathogenesis of disorders of the central nervous system. *Brain Pathol* 1993; 3: 77–85.
52. Terry RD, Masliah E, Salmon DP, et al. Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. *Ann Neurol* 1991; 30: 572–580.
53. Masliah E, Terry RD, DeTeresa RM, Hansen LA. Immunohistochemical quantification of the synapse-related protein synaptophysin in Alzheimer disease. *Neurosci Lett* 1989; 103: 234–239.
54. Masliah E, Mallory M, Hansen L, et al. Quantitative synaptic alterations in the human neocortex during normal aging. *Neurology* 1993; 43: 192–197.
55. Scheff SW, Price DA, Sparks DL. Quantitative assessment of possible age-related change in synaptic numbers in the human frontal cortex. *Neurobiol Aging* 2001; 22: 355–365.
56. Scheff SW, Price DA. Synaptic density in the inner molecular layer of the hippocampal dentate gyrus in Alzheimer disease. *J Neuropathol Exp Neurol* 1998; 57: 1146–1153.
57. DeKosky ST, Scheff SW, Styren SD. Structural correlates of cognition in dementia: quantification and assessment of synapse change. *Neurodegeneration* 1996; 5: 417–421.
58. Scheff SW, Price DA. Synapse loss in the temporal lobe in Alzheimer's disease. *Ann Neurol* 1993; 33: 190–199.
59. Scheff SW, Sparks L, Price DA. Quantitative assessment of synaptic density in the entorhinal cortex in Alzheimer's disease. *Ann Neurol* 1993; 34: 356–361.
60. McKeith IG, Perry EK, Perry RH. Report of the second dementia with Lewy body international workshop: diagnosis and treatment. Consortium on dementia with Lewy bodies. *Neurology* 1999; 53: 902–905.
61. McKeith IG, Dickson DW, Lowe J, et al. Diagnosis and management of dementia with Lewy bodies. Third report of the DLB consortium. *Neurology* 2005; 65: 1863–1872.

Amyloid-β Oligomers: Possible Roles as Key Neurotoxins in Alzheimer's Disease

Alex L. Lublin, PhD,[1] and Sam Gandy, MD, PhD[1,2]

[1]Department of Neurology, Mount Sinai Alzheimer's Disease Research Center,
Mount Sinai School of Medicine, New York, NY
[2]Department of Psychiatry, Mount Sinai Alzheimer's Disease Research Center,
Mount Sinai School of Medicine, New York, NY

OUTLINE

Amyloid Precursor Protein Processing
Oligomeric Amyloid-β Peptide in Alzheimer's Disease
Physiological Roles for Amyloid-β Peptide Oligomers
Conclusions

ABSTRACT

Alzheimer's disease is the most common form of senile dementia. Although the amyloid-β peptide was identified in 1984 as the major constituent of the senile plaques that characterize the disease, accumulating evidence indicates that the plaque density does not correspond well to the concurrent disease state. In order to resolve this disconnect, a number of recent studies have shifted away from the senile plaque and classical fibrillar forms of amyloid toward a less well structured species as the proximate neurotoxic factor underlying cognitive failure in Alzheimer's disease: soluble amyloid-β peptide oligomer (also known as the amyloid-β peptide–derived diffusible ligand). Paradoxically, several studies in the last 2 years have shown that picomolar levels of amyloid-β peptide have neutral activity or perhaps even an essential role in learning and memory. Here we highlight some of the key observations underlying the growing focus on the amyloid-β peptide oligomer. ***Mt Sinai J Med 77:43–49, 2010.*** © *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Sam Gandy
Mount Sinai Alzheimer's
Disease Research Center
Mount Sinai School of Medicine
New York, NY
Email: samuel.gandy@mssm.edu

Key Words: Alzheimer's disease, amyloid-β peptide, dementia, oligomer.

Alzheimer's disease (AD) is the most common cause of senile dementia. Clinically, AD patients exhibit progressive cognitive failure, including a loss of the ability to form and retrieve new memories, changes in personality, and a loss of the ability to navigate even the most familiar environments. Ultimately, all cortical function is lost, and death occurs as a complication of the terminal bed-bound state (ie, pneumonia or sepsis).

Pathologically, AD is defined by a characteristic loss of hippocampal and cerebrocortical neurons, especially those involving the diffuse cholinergic projection from the basal forebrain to the cerebral cortex. The neurodegenerative process in AD is also characterized by the accumulation of structural pathology in the form of interstitial and cerebrovascular deposits of amyloid and intraneuronal neurofibrillary tangles due to collapse of the cytoskeleton.

The neurodegenerative process in Alzheimer's disease (AD) is also characterized by the accumulation of structural pathology in the form of interstitial and cerebrovascular deposits of amyloid and intraneuronal neurofibrillary tangles due to collapse of the cytoskeleton.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20160

© 2010 Mount Sinai School of Medicine

The cerebral amyloid in AD, unlike that of other organs, is deposited not as a mass of fibrils but rather as miliary structures known as plaques, which primarily consist of the amyloid-β peptide (Aβ). During proteolytic processing of the Aβ parent or amyloid precursor protein (APP), the Aβ domain is excised by successive cleavage of APP by enzymes known as β-secretase and γ-secretase (discussed in more detail later). γ-Secretase cleavage generates multiple Aβ species, including soluble monomeric peptides of 40 amino acids and more insoluble peptides of 42 amino acids.

In AD, Aβ42 monomers will aggregate into progressively larger polymers (dimers, trimers, etc.) generically known as oligomers. A popular current concept regarding the pathogenesis of cognitive failure in AD is that Aβ oligomers are especially important neurotoxic conformers of Aβ.[1] This formulation contrasts with the traditional concept in which the key toxins are the highly structured Aβ fibrils that compose amyloid plaques. Furthermore, the current formulation is that the oligomer pathway and the amyloid fibril pathway are separate and distinct.

Initially, clinicopathological correlation studies brought the role of the amyloid plaque into question,[2–5] and more recently, the advent of amyloid imaging with Pittsburgh compound B has confirmed that many healthy adults accumulate numerous Aβ deposits with few signs of dementia.[2–5] The most compelling lines of evidence linking Aβ to AD are genetic because rare mutations either in APP or in 1 of 2 enzymes that act upon it (γ-secretase and α-secretase) can cause hereditary forms of AD that are clinically and pathologically indistinguishable from the more common, so-called sporadic form of the disease. These mutations apparently exert their actions by modulating APP processing (for a review, see Gandy[6]).

AMYLOID PRECURSOR PROTEIN PROCESSING

The formation of Aβ occurs though the sequential proteolysis of APP. APP is a type I, single-pass transmembrane protein that contains 3 potential sites for cleavage by various proteinases, which are designated α, β, and γ. α-Secretase cleavage by a disintegrin and metalloproteinase 10 (ADAM10) and ADAM17/transforming growth factor α–converting enzyme (or tumor necrosis factor α–converting enzyme), within the Aβ domain of APP, is the predominant and nonamyloidogenic pathway producing soluble amyloid precursor protein α (sAPPα) and the carboxyl terminal fragment C83. The β- and γ-cleavage pathway of APP produces sAPPβ and Aβ through successive cleavage by β-secretase (also known as BACE for β-amyloid precursor protein site cleaving enzyme) and the γ-secretase complex (Figure 1). The γ-secretase complex is made up of several substituents (presenilin 1, presenilin 2, presenilin enhancer 2, anterior pharynx-defective 1, and nicastrin).[7] All 4 proteins must be present in order for catalysis to proceed.

The amyloidogenic pathway produces Aβ of several different lengths ranging from 37 to 43 amino acids. The predominantly produced species is Aβ40. A trace amount of Aβ42 is produced in a ratio of approximately 99 to 1. Aβ42 spontaneously aggregates, and this makes it more likely to form oligomers; it is consequently considered to be the more toxic species. In familial Alzheimer's disease (FAD), there are mutations in APP, presenilin 1, presenilin 2, or ADAM10. Data gathered from inherited forms of AD show that genetic defects in any of these genes are associated with the enhanced accumulation of the more toxic Aβ42 species. Some mutations cause quantitative or qualitative changes in the cleavages that generate Aβ. Other mutations are localized within the Aβ sequence of APP and cause the generation of Aβ peptides that bear an enhanced inherent predisposition toward oligomerization,[8] and this provides further support for the oligomer hypothesis of Aβ neurotoxicity.

OLIGOMERIC AMYLOID-β PEPTIDE IN ALZHEIMER'S DISEASE

Several approaches have converged to implicate soluble oligomeric Aβ in AD. One such approach has employed a well-known electrophysiological correlate to learning and memory known as hippocampal long-term potentiation (LTP). Several studies have shown that hippocampal LTP can be inhibited by both synthetic and naturally secreted human Aβ oligomers. Some investigators have reported that the

Several studies have shown that hippocampal long-term potentiation, which is correlated with learning and memory, can be inhibited by both synthetic and naturally secreted human amyloid beta oligomers.

DOI:10.1002/MSJ

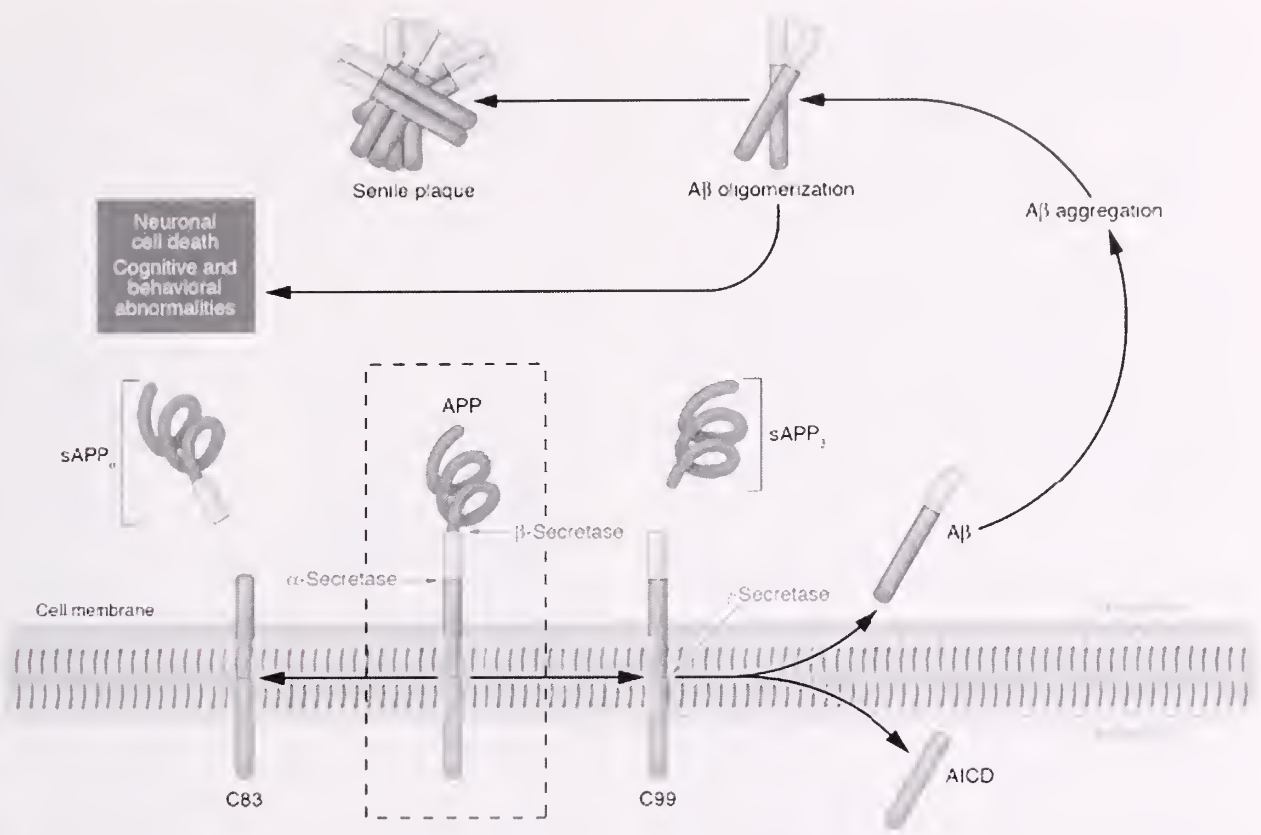


Fig 1. APP processing and Aβ accumulation. Mature APP (center and inside the dashed box) is metabolized by 2 competing pathways: the α-secretase pathway, which generates sAPPα and C83 (also known as C-terminal fragment α; left), and the β-secretase pathway, which generates sAPPβ and C99 (right). Some β-secretase cleavage is displaced by 10–amino acid residues and generates sAPPβ′ and C89. All carboxy-terminal fragments (C83, C99, and C89) are substrates for γ-secretase, which generates AICD, and the secreted peptides p3 (not shown), Aβ (right), and Glu11 Aβ. Aβ aggregates into small multimers (dimers, trimers, etc.) known as oligomers. Oligomers appear to be the most potent neurotoxins, whereas the end-stage senile plaque is relatively inert. **Abbreviations:** Aβ, amyloid-β peptide; AICD, amyloid precursor protein intracellular domain; APP, amyloid precursor protein; sAPP, soluble amyloid precursor protein. Reprinted with permission from *Journal of Clinical Investigation*.[6] Copyright 2005, American Society for Clinical Investigation.

oligomeric Aβ levels needed to disrupt LTP correlate well with the amount observed in cerebrospinal fluid in AD patients. Furthermore, the effects of Aβ oligomers on LTP can be reduced through the application of anti-Aβ antibodies in vivo.[9–12]

Although LTP correlates with enhanced learning and memory, the opposite can be said for long-term synaptic depression (LTD). Therefore, LTD in the hippocampus is a negative regulator of learning and memory. Using Aβ oligomers from several sources (cell cultures and synthetic and human brain extracts) on mouse hippocampal slices, Li *et al.*[11] demonstrated that oligomeric Aβ enhanced LTD. The enhancement of LTD could be blocked by the addition of a glutamate scavenger, and this suggests that oligomeric Aβ might facilitate LTD through a glutamate-dependent system.[11] Thus, oligomeric Aβ may lead to a disruption in learning and memory by inhibiting prosynaptic LTP and enhancing the silencing of LTD, and this appears to occur via a glutamatergic mechanism.

In addition to LTP, behavioral studies have been performed in which human Aβ oligomers were infused into the hippocampus of living wild-type rats. The direct application of oligomeric Aβ allows for the immediate study of the effects of acute treatment. In one study by Cleary *et al.*,[12] human Aβ oligomers were generated in Chinese hamster ovary cells and secreted into a conditioned medium from which these Aβ oligomers were collected. Rats were studied before and after treatment with intrahippocampal injections of Aβ and then placed in the Morris water maze, in which they had been pretrained to find a hidden platform in a pool of opaque liquid. These animals developed a deficit in learning and memory that was readily observable in the Morris water maze after the administration of Aβ42. This deficit was reversible upon washout or the application of

Rats developed a deficit in learning and memory that was readily observable in the Morris water maze after the administration of $A\beta42$.

an anti-$A\beta$ antibody. Thus, the in vitro LTP and LDP results appeared to be relevant in vivo as well.

Other brain regions are also involved in transgenic mouse models of AD. In a recent study, España *et al.*,[13] using the inhibitory avoidance/fear conditioning paradigm, found an enhanced level of freezing behavior in APP Indiana, APP Sweden/Indiana, and triple mutant (3xTg-AD) mouse models of AD. Using immunocytochemistry, they observed that these mutant mice had both extracellular amyloid deposits and intraneuronal $A\beta$-like immunoreactivity in neurons of the basolateral amygdala, a region of the brain known to modulate the acquisition of fear and fear-related behaviors. The observed increase in freezing behaviors by the mutant mice was ameliorated by the administration of several anxiolytics, and this suggests that the behavioral abnormalities were caused by an increase in anxiety. These findings suggest that $A\beta$ accumulation in the basolateral amygdala contributes to an increase in anxiety in the mouse model. The data also raise the possibility that intraneuronal $A\beta$ may play a role in $A\beta$ toxicity; this notion has been raised in the past and remains unresolved.

PHYSIOLOGICAL ROLES FOR AMYLOID-β PEPTIDE OLIGOMERS

Although it is clear that high levels of oligomeric $A\beta$ in mammalian models impair LTP, new evidence is emerging that $A\beta$ oligomers in normal physiological concentrations can positively regulate learning and/or memory. Garcia-Osta and Alberini[14] performed a series of inhibitory avoidance/fear conditioning experiments in which a rat was trained by the application of a painful foot shock in association with a given area of the test apparatus. The animal's aversion was then measured either 1 hour after training (short-term memory) or 24 hours after training (long-term memory). In one experiment, a monoclonal antibody specific to $A\beta$ (4G8) or a control antibody was administered 15 minutes prior to training. The results demonstrated a deficit in both the short-term and long-term memory tests after treatment with anti-$A\beta$ but not after treatment with the control antibody (Figure 2).[14] Alternatively, when picomolar amounts of exogenous human $A\beta42$ were added in combination with the anti-$A\beta$ antibody, a significant rescue of the amnesia was observed. In this instance, the interpretation is that the anti-$A\beta$ antibody bound endogenous $A\beta$ that was putatively involved in physiological learning and memory processing.

New evidence is emerging that $A\beta$ oligomers in normal physiological concentrations can positively regulate learning and/or memory.

Similar experiments were performed on mice by Puzzo *et al.*[15] Using an experimental system similar to the one that Cleary *et al.*[12] employed to demonstrate the detrimental effects of high levels of oligomeric $A\beta$, Puzzo *et al.* were able to show a learning and memory benefit when picomolar quantities of oligomeric $A\beta$ were used. Then, using brain sections, these same investigators demonstrated that LTP was increased in mice that received picomolar quantities of $A\beta$, but when levels of $A\beta$ over 200 pM were applied, a deficit developed. This was reinforced by behavioral studies in which those mice that were treated with picomolar $A\beta$ had an enhanced Morris water maze performance and better fear memory in inhibitory avoidance/fear conditioning tests.[15] These observations suggest that oligomeric $A\beta$ has an inverted U-shaped dose-response curve, which is consistent with the proposal that this peptide has mnemonoactive properties.

In addition to the $A\beta$ concentration, however, current data suggest that the different conformations and species of oligomeric $A\beta$ are functionally distinct. To address this issue, Murray *et al.*[16] used mass spectrometry coupled with ion mobility spectrometry to determine the number of $A\beta$ peptides per aggregate in a series of experimental solutions.

Solutions containing $A\beta40$ only, $A\beta42$ only, or a mixture of the two were analyzed for the relative quantities of the differently sized oligomers (ie, dimers, tetramers, hexamers, etc.). $A\beta40$ oligomers reached a maximum size at the tetramer level, whereas $A\beta42$ was capable of producing much larger oligomers (ie, up to dodecamers; Figure 3). However, when $A\beta40$ and $A\beta42$ were combined in solution, the largest resulting oligomer was only a tetramer. Therefore, $A\beta40$ is capable of suppressing the formation of the larger oligomers seen with $A\beta42$ alone.[16] In addition to suggesting new pharmaceutical targets, this finding adds strength to the theory that the ratio of $A\beta40$ to $A\beta42$ is a key determinant of risk for AD.

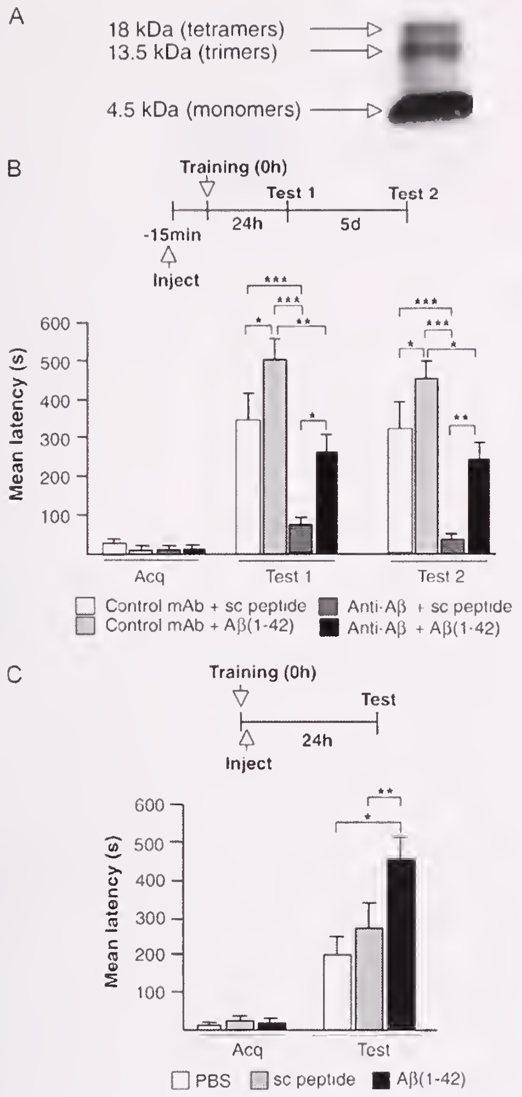


Fig 2. Memory impairment produced by the depletion of endogenous Aβ(1–42) is rescued by exogenous oligomeric human Aβ(1–42). (A) An oligomeric/monomeric preparation of Aβ42 was examined via 4% to 12% tris-tricine nondenaturing polyacrylamide gel electrophoresis/western blotting with the anti-Aβ monoclonal antibody 6E10 (Covance Research; 1:1000). Bands corresponding to tetramers, trimers, and monomers were detected. (B) Acq and retention, expressed as the mean latency ± SEM (seconds), of rats that received intrahippocampal injections of either anti-Aβ or a control monoclonal antibody combined with either sc peptide or Aβ(1–42) 15 minutes before IA training. $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. Test 1 occurred 24 hours after training; test 2 occurred 5 days after training. (C) Acq and retention, expressed as the mean latency ± SEM (seconds), of rats that received intrahippocampal injections of PBS, sc peptide, or Aβ(1–42) immediately after IA training. The administration of Aβ(1–42) immediately after IA training enhanced memory retention 24 hours after training. $^{*}P < 0.05$ and $^{**}P < 0.01$. **Abbreviations:** Aβ, amyloid-β peptide; Acq, memory acquisition; IA, inhibitory avoidance; PBS, phosphate-buffered saline; sc, scrambled; SEM, standard error of the mean. Reprinted with permission from *Learning & Memory*.[14] Copyright 2009, Cold Spring Harbor Press.

Another possible explanation for the Aβ benefit/deficit paradox invokes a role for monomeric Aβ. Guiffida *et al.*[17] demonstrated that cultured neurons, under conditions of insulin deprivation, appeared to benefit when monomeric Aβ42 was administered. This benefit correlated with an observed increase in the phosphorylation of Akt, and this suggests that the rescue of the neurons occurred through the insulin-responsive phosphatidylinositol-3-kinase pathway. Monomers of Aβ42 were also able to protect against an *N*-methyl-d-aspartic acid–induced excitotoxic death, the opposite of which is observed in the presence of aggregated Aβ. Interestingly, there was no neuronal protection when the monomeric Aβ42 contained the Arctic (E22G) substitution, a rare FAD mutation that leads to rapid Aβ oligomerization. These results suggest that a lack of monomeric Aβ can be detrimental. Therefore, the increased formation of oligomers that occurs in association with many FAD mutations might cause neurotoxicity, at least in part by depleting the supply of beneficial Aβ monomers while simultaneously increasing the levels of potentially neurotoxic Aβ oligomers.[17]

Another recent study sheds more light on the toxicity of Aβ42. In the amino acid sequence of Aβ42, there is a hydrophobic GxxxG motif that is vital for dimerization. Substitution of Gly33 for the more hydrophobic residues of either alanine or isoleucine within the dimerization motif resulted in a marked increase in the formation of higher order oligomers (16- to 20-mers). Surprisingly, even with larger aggregates, the Gly33 substitution proved to be nontoxic in comparison with normal Aβ42 in both neuronal cells and the *Drosophila* eye photoreceptor assay. Furthermore, it was found that substitution of Gly33 abolished the deleterious effect that Aβ42 has on LTP, and this indicates that Gly33 is necessary for Aβ42 toxicity and, more importantly, that toxicity is not directly correlated with all higher order oligomers because dimers, trimers, and tetramers were the most toxic in this experimental paradigm.[18]

The issue of what happens when extracellular oligomeric Aβ binds to neurons was recently addressed by Laurén *et al.*[19] These researchers suggested that cellular prion protein (PrPc) might act as a neuronal receptor for oligomeric Aβ. As previously shown, oligomeric Aβ inhibits LTP in the hippocampus of wild-type mice. In Laurén *et al.*'s study, mice lacking PrPc exhibited little or no inhibition of LTP after treatment with oligomeric Aβ. Furthermore, blocking PrPc with antibodies was also sufficient to ameliorate the detrimental effects of oligomeric Aβ. However, PrPc accounted for only 50% of the total neuronal binding of oligomeric Aβ.[19] Additional binding of Aβ is presumably due to other

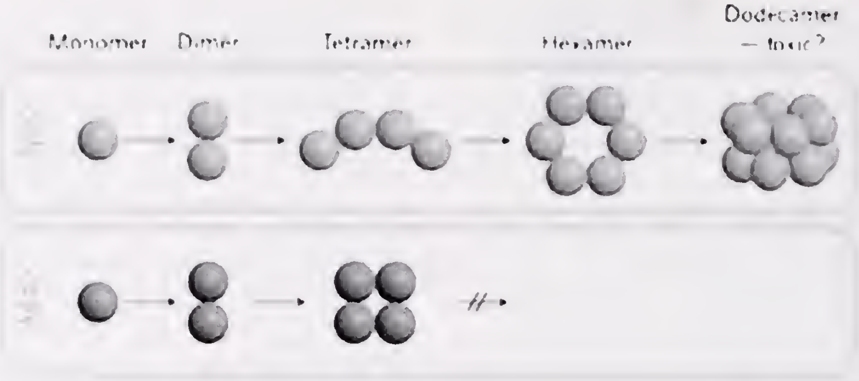


Fig 3. Relationship linking the Aβ peptide species to the oligomerization state and neurotoxicity. **Abbreviation:** Aβ, amyloid-β peptide. Reprinted with permission from *Nature Chemistry*.[20] Copyright 2009, Nature Publishing Group.

receptors, perhaps including low-density lipoprotein receptors, which have been shown to bind several ligands, including Aβ. Alternatively, there could be as yet unknown factors that are responsible for binding oligomeric Aβ. More studies are required to confirm and elucidate this potential role for PrP in Aβ neurotoxicity.

CONCLUSIONS

Recent advances initially demonstrated that soluble oligomers of Aβ may be key to the pathogenesis of AD. However, with the dissection of the concentration dependence and assembly state dependence, soluble oligomeric Aβ species are now believed to play vital physiological functions in learning and memory. The discovery of PrPc as a functional receptor of Aβ oligomers may enhance our understanding of how Aβ leads to cognitive defects. However, the binding of Aβ by PrPc is responsible for only 50% of the observed neuronal oligomer binding capacity, and this raises the question of what other receptors for oligomeric Aβ might exist. The evolving concepts of amyloid plaques, amyloid fibrils, and amyloid oligomers point to the necessity of reflection and continual re-evaluation of our concepts of physiology and pathogenesis in normal brain function and in AD.

ACKNOWLEDGMENT

This work was supported by NIA P01 AG10491 (to Sam Gandy), P50 AG05138 (to Mary Sano), and the Cure Alzheimer's Fund.

DISCLOSURES

Potential conflict of interest: Nothing to report.

DOI:10.1002/MSJ

REFERENCES

1. Lambert MP, Barlow AK, Chromy BA, et al. Diffusible, nonfibrillar ligands derived from Abeta1–42 are potent central nervous system neurotoxins. *Proc Natl Acad Sci U S A* 1998; 95: 6448–6453.
2. Giannakopoulos P, Herrmann FR, Bussière T, et al. Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology* 2003; 60: 1495–1500.
3. Näslund J, Haroutunian V, Mohs R, et al. Correlation between elevated levels of amyloid beta-peptide in the brain and cognitive decline. *JAMA* 2000; 283: 1571–1577.
4. Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. *Ann Neurol* 1999; 45: 358–368.
5. Aizenstein HJ, Nebes RD, Saxton JA, et al. Frequent amyloid deposition without significant cognitive impairment among the elderly. *Arch Neurol* 2008; 65: 1509–1517.
6. Gandy S. The role of cerebral amyloid beta accumulation in common forms of Alzheimer disease. *J Clin Invest* 2005; 115: 1121–1129.
7. Haass C. Take five–BACE and the gamma-secretase quartet conduct Alzheimer's amyloid beta-peptide generation. *EMBO J* 2004; 23: 483–488.
8. Haass C, Selkoe DJ. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nat Rev Mol Cell Biol* 2007; 8: 101–112.
9. Selkoe DJ. Soluble oligomers of the amyloid beta-protein impair synaptic plasticity and behavior. *Behav Brain Res* 2008; 192: 106–113.
10. Shankar GM, Bloodgood BL, Townsend M, et al. Natural oligomers of the Alzheimer amyloid-beta protein induce reversible synapse loss by modulating an NMDA-type glutamate receptor-dependent signaling pathway. *J Neurosci* 2007; 27: 2866–2875.
11. Li S, Hong S, Shepardson NE, et al. Soluble oligomers of amyloid β protein facilitate hippocampal long-term depression by disrupting neuronal glutamate uptake. *Neuron* 2009; 62: 788–801.
12. Cleary JP, Walsh DM, Hofmeister JJ, et al. Natural oligomers of the amyloid-beta protein specifically disrupt cognitive function. *Nat Neurosci* 2005; 8: 79–84.

13. España J, Giménez-Llort L, Valero J, et al. Intraneuronal beta-amyloid accumulation in the amygdala enhances fear and anxiety in Alzheimer's disease transgenic mice. *Biol Psychiatry*. doi://10.1016/j.biopsych.2009.06.015. Available at: http://www.sciencedirect.com.
14. Garcia-Osta A, Alberini CM. Amyloid beta mediates memory formation. *Learn Mem* 2009; 16: 267–272.
15. Puzzo D, Privitera L, Leznik E, et al. Picomolar amyloid-beta positively modulates synaptic plasticity and memory in hippocampus. *J Neurosci* 2008; 28: 14537–14545.
16. Murray MM, Bernstein SL, Nyugen V, et al. Amyloid beta protein: Abeta40 inhibits Abeta42 oligomerization. *J Am Chem Soc* 2009; 131: 6316–6317.
17. Giuffrida ML, Caraci F, Pignataro B, et al. β-Amyloid monomers are neuroprotective. *J Neurosci* 2009; 29: 10582–10587.
18. Harmeier A, Wozny C, Rost BR, et al. Role of amyloid-beta glycine 33 in oligomerization, toxicity, and neuronal plasticity. *J Neurosci* 2009; 29: 7582–7590.
19. Laurén J, Gimbel DA, Nygaard HB, et al. Cellular prion protein mediates impairment of synaptic plasticity by amyloid-beta oligomers. *Nature* 2009; 457: 1128–1132.
20. Bernstein SL, Dupuis NF, Lazo ND, et al. Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nat Chem* 2009; 1: 326–331.

Relationship Between Tau Pathology and Neuroinflammation in Alzheimer's Disease

Maria Jose Metcalfe, MS, and Maria E. Figueiredo-Pereira, PhD

Department of Biological Sciences, Hunter College, City University of New York, New York, NY

OUTLINE

Tau Phosphorylation
Protein Interacting with Nima (Never in Mitosis gene A) 1 Restores Tau Function
Tau Cleavage
Correlation Between Tau Hyperphosphorylation and Caspase Cleavage
Tau Ubiquitination
Role in Tau Degradation of Heat Shock Protein 70- Interacting Protein/Heat Shock Protein 90 Complex
Inflammation and Tau Pathology
Late-Breaking Studies
Conclusion

ABSTRACT

Alzheimer's disease is a chronic, age-related neurodegenerative disorder. Neurofibrillary tangles are among the pathological hallmarks of Alzheimer's disease. Neurofibrillary tangles consist of abnormal protein fibers known as paired helical filaments. The accumulation of paired helical filaments is one of the most characteristic cellular changes in Alzheimer's disease. Tau protein, a microtubule-associated protein, is the major component of paired helical filaments. Tau in paired helical filaments is hyperphosphorylated, truncated, and aggregated. What triggers the formation of paired helical filaments is not known, but neuroinflammation could play a role. Neuroinflammation is an active process detectable in the earliest stages of Alzheimer's disease. The neuronal toxicity associated with inflammation makes it a potential risk factor in the pathogenesis of Alzheimer's disease. Determining the sequence of events that lead to this devastating disease has become one of the most important goals for the prevention and treatment of Alzheimer's disease. In this review, we focus on the pathological properties of tau thought to play a role in neurofibrillary tangle formation and summarize how central nervous system inflammation might be a critical contributor to the pathology of Alzheimer's disease. A better understanding of the mechanisms that cause neurofibrillary tangle formation is of clinical importance for developing therapeutic strategies to prevent and treat Alzheimer's disease. One of the major challenges facing us is singling out neuroinflammation as a therapeutic target for the prevention of Alzheimer's disease neurodegeneration. The challenge is developing therapeutic strategies that prevent neurotoxicity linked to inflammation without compromising its neuroprotective role. ***Mt Sinai J Med 77:50–58, 2010.*** *© 2010 Mount Sinai School of Medicine*

Address Correspondence to:

Maria E. Figueiredo-Pereira
Department of Biological
Sciences Hunter College City
University of New York
New York, NY
Email: pereira@genectr.hunter.cuny.edu

Key Words: aggregation, Alzheimer's disease, inflammation, neurofibrillary tangles, tau protein.

Alzheimer's disease (AD) is histopathologically characterized by extracellular β-amyloid–containing plaques and intracellular tau-containing neurofibrillary tangles (NFTs).[1] The main component of NFTs is tau, a highly soluble microtubule-associated protein whose major function is to stabilize microtubules, specifically in axons, in a phosphorylation-dependent manner.[2] Tau promotes the assembly of microtubules

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20163

The main component of neurofibrillary tangles (NFTs) is tau, a highly soluble microtubule-associated protein whose major function is to stabilize microtubules, specifically in axons, in a phosphorylation-dependent manner.

into evenly spaced bundles in the axons and regulates the growing and shortening dynamics of individual microtubules.[3] Tau is a highly soluble protein that is heat- and acid-stable, and it is rich in polar and charged amino acids with a basic character.[2] Five residues (glycine, lysine, proline, serine, and threonine) make up half of the tau sequence, and this justifies its high solubility and the unfolded nature of the protein.

The tau protein has 4 different regions (Figure 1), which range from the C-terminus to the N-terminus:

1. The C-terminal domain with acidic and basic subregions, which indirectly control tau binding to microtubules via regulated phosphorylation.
2. The microtubule-binding domain with 3 or 4 imperfect repeats.
3. The positively charged proline-rich domain, which indirectly regulates the association between tau and microtubules via regulated phosphorylation.
4. The N-terminal domain, which contains 0, 1, or 2 negatively charged inserts.[2]

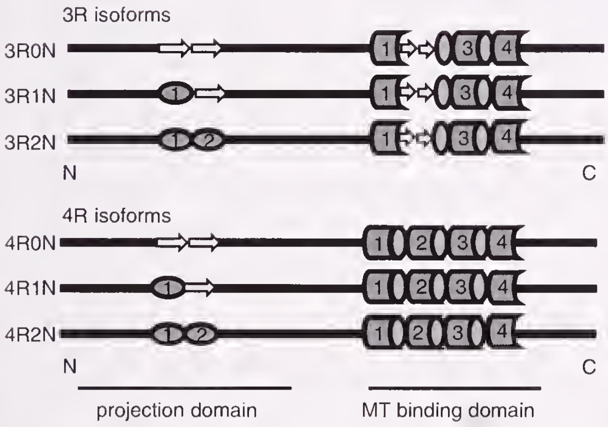


Fig 1. Schematic representation of the 6 tau isoforms generated by alternative RNA splicing. The N-terminus contains zero, one, or two 29-aa inserts (pink) and, together with the proline-rich domain, forms the projection domain, which protrudes away from MTs. The C-terminus includes either three or four 18-aa-long imperfect repeats (blue) separated by 13- to 14-aa-long inter-repeats (green), and together they form the MT-binding domain. **Abbreviations:** aa, amino acid; MT, microtubule.

The proline-rich domain and the N-terminus compose the projection domain, which extends outward from the microtubule surface.[3]

There are 6 tau isoforms in the human brain, and they are products of alternative RNA splicing at exons 2, 3, and 10.[4] The 6 different tau isoforms can be designated 3R0N, 3R1N, 3R2N, 4R0N, 4R1N, and 4R2N, the designation depending on the number of N-terminal inserts and C-terminal repeats. The microtubule-binding region contains 3 (3R) or 4 (4R) imperfect 18-amino acid-long repeats with a characteristic Pro-Gly-Gly-Gly motif separated from one another by an inter-repeat that is 13 to 14 amino acids long.[2,5] The fetal human brain expresses only 3R tau, whereas the adult brain expresses approximately equal amounts of 3R tau and 4R tau. Mutations in the tau gene lead to extensive neuronal cell death and dementia, which is manifested in diseases such as hereditary Pick's disease and frontotemporal dementia with parkinsonism linked to chromosome 17.[3] Exonic mutations involve amino acid substitutions in the microtubule-binding repeat region or close to it, and this suggests that they impair tau binding to microtubules. Intronic mutations located near the splice-donor site of the intron following exon 10 are regulatory and affect tau RNA splicing without affecting the amino acid composition of the protein.[6]

Neurodegenerative diseases collectively designated *tauopathies* are linked to tau mutations and/or tau posttranslational modifications. Accordingly, tau hyperphosphorylation and cleavage are important events leading to tau intracellular accumulation, aggregation, and neuronal cell death.[7]

Neurodegenerative diseases collectively designated tauopathies are linked to tau mutations and/ or tau posttranslational modifications.

TAU PHOSPHORYLATION

An important characteristic of tau is that it possesses a large number of potential phosphorylation sites. Kinases that phosphorylate tau are distributed into 2 groups[8]:

1. Proline-directed serine-threonine protein kinases that phosphorylate Ser-Pro or Thr-Pro tau motifs. These include glycogen synthase kinase 3β

(GSK3β), cyclin-dependent kinase 5 (cdk5), and stress kinases such as c-Jun N-terminal kinase (JNK) and p38.

2. Non–proline-directed protein kinases that phosphorylate serine or threonine residues not followed by proline. These include protein kinases A and C, calcium calmodulin–dependent kinase II, serum and glucocorticoid-dependent kinase, protein kinase B (Akt), microtubule-affinity regulating kinase, and synapses of amphids defective kinases, which are also known as brain-specific serine/threonine kinases.

The phosphorylation of tau to physiological levels is a way in which tau function is regulated.[9] The physiological role of tau phosphorylation includes regulation of microtubule dynamics, neurite outgrowth, and axonal transport.[9] Under normal physiological conditions, there is a balance between tau phosphorylation and dephosphorylation. Phosphatases such as protein phosphatase 1 (PP1), PP2A, PP2B, and PP2C are known to reverse tau phosphorylation, with PP2A postulated to be the phosphatase that dephosphorylates most tau phosphorylation sites.[10] Abnormal hyperphosphorylation of tau within a consensus sequence for proline-directed kinases is an important factor in the conversion of normal tau into paired helical filaments (PHFs).[3] For example, combinations of the kinases cdk5/GSK3 and calcium calmodulin kinase II/GSK3β are involved in the rapid phosphorylation of tau at Thr231 and Ser235, which is required for PHF formation in AD.[11,12]

Dephosphorylation of tau by PP2A inhibits its aggregation into PHFs and restores its ability to bind to microtubules. However, rephosphorylation of tau by different combinations of protein kinase A, calcium calmodulin kinase II, GSK3β, and cdk5 promotes its self-assembly into PHFs, and this leads to NFT formation.[12]

When tau is abnormally hyperphosphorylated, it loses its biological activity, becomes resistant to degradation, and goes through conformational changes that render it insoluble and aggregation-prone. Abnormal hyperphosphorylation of tau leads to its aggregation into PHFs, which are the main components of NFTs. Understanding PHF assembly and how to prevent it is relevant to AD and other neurodegenerative diseases in which tau plays a critical role, such as frontotemporal dementia with parkinsonism linked to chromosome 17.

When tau is abnormally hyperphosphorylated, it loses its biological activity, becomes resistant to degradation, and goes through conformational changes that render it insoluble and aggregation-prone.

PROTEIN INTERACTING WITH NIMA (NEVER IN MITOSIS GENE A) 1 RESTORES TAU FUNCTION

The large family of proline-directed serine-threonine protein kinases, which include GSK3β, cdk5, and stress kinases such as JNK, plays a critical role in cellular growth regulation, stress responses, and neuronal survival.[13] These kinases specifically recognize serine and threonine residues that precede proline. Because of its unique stereochemistry, proline can adopt 2 different conformational states: cis-isomeric and trans-isomeric states.[13] Although the uncatalyzed isomerization of proline is slow, it can be accelerated by peptidyl-prolyl cis/trans-isomerases, which are enzymes able to change protein conformation. One of these enzymes, a peptidyl-prolyl isomerase called protein interacting with NIMA (never in mitosis gene A) 1 (PIN-1), regulates phosphorylation signaling by proline-directed serine-threonine protein kinases.[13]

The peptidyl-prolyl isomerase PIN-1 has been linked to AD because it has been detected in NFTs from AD brains[14] and is thought to be neuroprotective against AD neurodegeneration.[15] PIN-1 possesses an N-terminal WW domain responsible for mediating the binding of PIN-1 to tau, specifically at the pT231 residue. PIN-1 promotes tau dephosphorylation at the pThr231-Pro motif by the trans-specific phosphatase PP2A. Notably, PIN-1 binds to tau only after its mitosis-specific phosphorylation.[14] Microtubule polymerization assays in conjunction with PIN-1 localization studies in normal and AD brains have suggested that one of the functions of PIN-1 is to bind directly to pThr231 on tau, promote its dephosphorylation, and restore the ability of tau to bind to microtubules.[14]

The peptidyl-prolyl isomerase that interacts with never-in-mitosis-A 1 has been linked to Alzheimer's disease (AD) because it has been detected in NFTs from AD brains and is thought to be neuroprotective against AD neurodegeneration.

Neurons are postmitotic cells, but some studies support the view that aberrant neuronal cell cycle re-entry plays an important role in AD neurodegeneration.[16–18] Aberrant neuronal reactivation of mitosis is associated with tau hyperphosphorylation. A reduction of PIN-1 protein levels, as observed in AD, will result in higher levels of hyperphosphorylated tau; this compromises its ability to promote microtubule polymerization and increases tau aggregation and NFT formation. PIN-1 and the associated prolyl isomerization process are thus potential therapeutic targets for preventing AD neurodegeneration.[19]

TAU CLEAVAGE

The independent cleavage of tau by caspases and calpain is an important posttranslational modification that, together with hyperphosphorylation, plays a critical role in the formation of PHFs.[12] Caspase-cleaved tau is detected in NFTs, and this supports the view that the apoptosis cascade is involved in the formation of NFTs.[20] It is thought that tau cleavage at its C-terminus by caspases renders tau prone to hyperphosphorylation and the formation of NFTs.[2] The N-terminal of tau is cleaved by calpain, which induces a change in the tau conformation from an unfolded state to a β-sheet structure that renders its C-terminus susceptible to cleavage by caspases.[12]

Caspase-cleaved tau is detected in NFTs, and this supports the view that the apoptosis cascade is involved in the formation of NFTs.

The role of apoptosis in AD neurodegeneration is still unclear, although caspases seem to be activated early in the progression of AD.[21] The 2 main apoptotic pathways are (1) the death receptor pathway involving the initiator caspase 8 and (2) the mitochondrial pathway involving the initiator caspase 9.[22] Both pathways converge on pro-caspase 3 cleavage, which leads to a cascade that triggers caspase 3 activation, tau cleavage, and tau pathological aggregation. In AD brains, a colocalization of caspase 8[23] and caspase 9[24] with NFTs has been observed, and this suggests that both apoptotic pathways occur within the same neuronal populations in AD brains. However, it is not clear whether caspase 8 or 9 activation precedes or coincides with NFT formation.[24]

Tau is cleaved by multiple caspases at a highly conserved aspartic acid residue (Asp421) in its C-terminus, and this produces the N-terminal product Asp421Tau.[20,25] Cleavage assays of full-length tau in the presence of various caspases have shown that tau is more susceptible to cleavage by executioner caspases (caspase-3 and caspase-7) than by initiator caspases (caspase-1, caspase-4, caspase-5, caspase-8, and caspase-10).[26] Tau aggregation assays have indicated that Asp421Tau, generated by caspase 3 or caspase 9, aggregates faster and has a stronger seeding effect for aggregation than full-length tau.[24]

In AD brains, Asp421Tau is widespread in the cornu ammonis 1 region of the hippocampus. This distribution colocalizes with intraneuronal and extraneuronal amyloid beta peptide (Aβ) deposits and correlates with cognitive impairment.[22] Asp421Tau has also been detected in other tauopathies, such as Pick's disease,[27] supranuclear palsy, cortical degeneration, and dementia with Lewy bodies.[28]

There seems to be a correlation between Aβ, the major component of senile plaques in AD, and NFT formation. Aβ may lead to the activation of apoptosis through the death receptor as well as the mitochondrial pathways. Studies with E18 rat primary cortical neurons have shown that upon treatment with Aβ, tau is cleaved to Asp421Tau.[22,24] Other studies have demonstrated that tau cleavage induced by Aβ treatment is prevented when the cultures are pre-incubated with caspase inhibitors.[20] Furthermore, the treatment of hippocampal neurons with Aβ induces neurite degeneration and microtubule collapse only when tau is present. Tau-depleted neurons show no signs of degeneration in the presence of Aβ, and this supports a role for tau in Aβ-induced neurodegeneration.[29]

CORRELATION BETWEEN TAU HYPERPHOSPHORYLATION AND CASPASE CLEAVAGE

The relationship between tau hyperphosphorylation and its cleavage by caspases remains poorly defined. Some studies have suggested that phosphorylation precedes cleavage in tangle evolution.[23] In vitro phosphorylation of tau at Ser422 renders tau more resistant to caspase 3 proteolysis, and this supports the notion that phosphorylation at Ser422 prevents caspase cleavage some time during the progression of AD.[23]

The JNK family is involved in processes such as cell differentiation, proliferation, apoptosis, and neurodegeneration.[30] JNKs are activated under stress conditions, such as those induced by reactive oxygen species and ultraviolet radiation.[31] Studies using cell culture models[32] have established that JNKs induce tau hyperphosphorylation leading to caspase activation and thus promote tau cleavage. The JNK signaling pathway can be activated by a number of stress factors, including oxidative stress and pro-inflammatory cytokines.[33] JNK pathways are altered in AD; this causes abnormal phosphorylation of proteins that, under normal homeostatic conditions, would not be JNK targets.[30] There are many potential substrates for JNK, but there is great interest in determining whether JNK activation is involved in tau phosphorylation and if this process occurs before or after caspase cleavage and tau aggregation. Tau phosphorylation by JNK primes tau for phosphorylation by GSK3β, and this results in tau hyperphosphorylation. Only then will tau form toxic aggregates that will in turn activate caspases and induce neuronal death. This sequence of events is further supported by the colocalization of phospho-JNK with tau inclusions.[34]

TAU UBIQUITINATION

Whether tau is ubiquitinated and degraded in vivo by the ubiquitin proteasome pathway remains controversial. In addition, it is not clear if the tau present in NFTs is ubiquitinated. However, it is well established that proteasome activity is impaired in AD brains.[35]

There are least 2 mechanisms associated with AD that could be responsible for proteasome impairment. The first mechanism is associated with a frameshift mutant of ubiquitin B (UbB+1) found to colocalize with NFTs and senile plaques in the cerebral cortex of patients with sporadic AD.[36] The UbB+1 mutant lacks the C-terminal glycine, which is a critical amino acid for ubiquitination and the formation of polyubiquitin chains. UbB+1 impairs the degradation of ubiquitinated proteins by competing with wild-type ubiquitin for binding to the 26S proteasome.[37] The second mechanism linked to AD is related to the direct binding of PHFs to the proteasome, an event that impairs proteasome activity.[38]

A direct link between tau aggregation and proteasome impairment is missing, although PHF-tau has been found to be ubiquitinated at its microtubule-binding domain, and this suggests that tau ubiquitination may be an early pathological event in the AD cascade.[39] Some in vitro and cell culture studies have suggested that tau is degraded by the proteasome.[40] Tau degradation by the proteasome seems to be dependent on tau ubiquitination with K63-polyubiquitin chains. These chains are required for tau interaction with sequestosome 1/p62, a shuttling partner for some proteasome substrates.[41] Other studies using rat primary hippocampal neurons[42] and human neuroblastoma SH-SY5Y cells[43] have reported that tau degradation is independent of proteasome activity. This controversy remains to be solved.

> *A direct link between tau aggregation and proteasome impairment is missing, although paired-helical-filament tau has been found to be ubiquitinated at its microtubule-binding domain, and this suggests that tau ubiquitination may be an early pathological event in the AD cascade.*

Similar to ubiquitination, sumolation is a post-translational modification consisting of the addition of small ubiquitin-like modifiers to lysine residues on target proteins. Protein sumolation is known to prevent proteasomal degradation and/or to change protein function.[44] Tau sumolation has been found to be up-regulated in cells treated with phosphatase inhibitors, and this suggests a new regulatory modification of tau that may have an implication for its pathology.[39]

ROLE IN TAU DEGRADATION OF HEAT SHOCK PROTEIN 70-INTERACTING PROTEIN/HEAT SHOCK PROTEIN 90 COMPLEX

Tau undergoes hyperphosphorylation before its aggregation. One of the hypotheses for the neuronal accumulation of fibrillar tau in neurodegenerative disorders exhibiting NFTs is that chaperones fail to prevent the aggregation of hyperphosphorylated tau. Hyperphosphorylated tau is recognized and ubiquitinated by the U-box protein called carboxyl terminus of heat shock protein 70–interacting protein (CHIP).[45] CHIP is a cochaperone with intrinsic ubiquitin ligase activity, which allows its chaperone function to switch from protein folding to protein degradation.[44] CHIP interacts with heat shock protein 70 (Hsp70) and inhibits its ATPase activity, and this suggests that it has another function besides protein refolding. CHIP also binds to Hsp90 and forms a complex that selectively ubiquitinates hyperphosphorylated tau along with Ubiquitin-conjugating enzyme H5B (UbcH5B).[45] The levels of CHIP are inversely proportional to insoluble tau accumulation in AD brains.[46] In addition, mice lacking CHIP exhibit high levels of insoluble tau in their brains.[47] These results suggest that CHIP complexed with Hsp90 delays the formation of tau aggregates and may thus play a protective role. The serine/threonine kinase Akt is another protein that is ubiquitinated by the CHIP/Hsp90 complex and then targeted for proteasomal degradation.[48] Akt is one of the kinases that hyperphosphorylates tau and promotes tau accumulation/aggregation.[49] Studies using a variety of cell models have indicated that under stress conditions the CHIP/Hsp90 complex ubiquitinates both Akt and tau for proteasomal degradation.[48] Low Akt levels cause down-regulation of CHIP expression either by direct transcriptional regulation or via a feedback mechanism.[48] However, with aging or under disease conditions, the levels of Akt increase, and because the CHIP/Hsp90 complex has a higher affinity for Akt than for hyperphosphorylated tau, Akt is preferentially ubiquitinated and degraded. As a result, hyperphosphorylated tau fails to be degraded, and its levels increase; this promotes tau accumulation/aggregation.[48,50] Higher levels of tau increase its chances of being hyperphosphorylated by kinases such as Akt and microtubule-affinity regulating kinase/Prader-Willi/Angelman region 1,[51] the latter being a microtubule-affinity regulating kinase that plays a role in phosphorylating tau during AD neurodegeneration. Notably, Prader-Willi/Angelman region 1 phosphorylation of tau at S262/S356 sharply decreases tau affinity for the CHIP/Hsp90 complex and, therefore, enhances tau accumulation.[52] In conclusion, the role played by Akt in tau accumulation supersedes its role in tau phosphorylation.

The carboxyl terminus of heat shock protein 70-interacting protein complexed with heat shock protein 90 delays the formation of tau aggregates and may thus play a protective role.

INFLAMMATION AND TAU PATHOLOGY

Inflammation is implicated in AD.[53] The association between inflammation and AD manifests itself through (1) the presence of activated microglia and astrocytes surrounding senile plaques and (2) higher levels of inflammatory mediators in the brains of AD patients versus age-matched controls.[54] Current animal models of AD fail to address the mechanisms by which products of inflammation produced by activated microglia and astrocytes induce AD neurodegeneration. This is a very important issue because some products of inflammation are neuroprotective and others are neurotoxic.

A recent study using P301S mutant human tau transgenic mice established that hippocampal synaptic pathology and microgliosis could be the earliest manifestations of neurodegeneration related to tauopathies.[55] Prominent microglial activation preceded tangle formation, and the immunosuppression of young P301S transgenic mice diminished tau pathology and increased their lifespan. It was concluded that neuroinflammation is linked to the early progression of tauopathies.[55]

What triggers the formation of NFTs is not known, but neuroinflammation could play a critical role. Neuroinflammation is an active process detectable in the earliest stages of AD. The neuronal toxicity associated with inflammation makes it a potential risk factor in the pathogenesis of chronic neurodegenerative diseases such as AD. Determining the sequence of events that lead to this devastating disease has become one of the most important goals for AD prevention and treatment. We recently reviewed the link between inflammation and tau pathology.[56]

The neuronal toxicity associated with inflammation makes it a potential risk factor in the pathogenesis of chronic neurodegenerative diseases such as AD.

We propose a model (Figure 2) in which any stimulus (physical, chemical, or infectious) capable of inducing inflammation in brain areas affected in AD directly activates microglia and astrocytes. These stimuli can also directly induce neuronal injury, which through yet unknown factors triggers the reactive response of glial cells. In either case, the activated glia surround and drench neurons with their toxic products, such as prostaglandins, nitric oxide, interleukin 1β, interleukin 6, tumor necrosis factor alpha, and reactive oxygen species (eg, superoxide anion). All these cytotoxic agents must work in concert to induce synergistic neurotoxicity leading to neurodegeneration. This sequence of events could explain many pathological features of the AD neurodegenerative

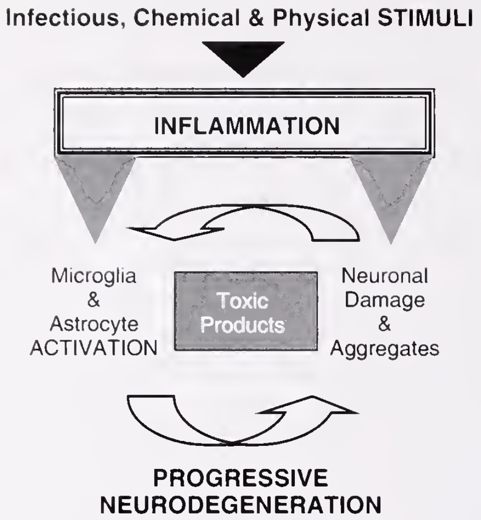

Infectious, Chemical & Physical STIMULI
INFLAMMATION
Microglia & Astrocyte ACTIVATION
Toxic Products
Neuronal Damage & Aggregates
PROGRESSIVE NEURODEGENERATION

Fig 2. Model for the relation between neuroinflammation and tau pathology in Alzheimer's disease. We propose a sequence of events initiated by stimuli (physical, chemical, or infectious) that induce inflammation. Activated glia (microglia and astrocytes) release inflammatory factors, such as some prostaglandins (eg, prostaglandin J2), nitric oxide, interleukin 6, tumor necrosis factor alpha, and reactive oxygen species (eg, superoxide anion), that work in concert to induce neurotoxicity. One of the consequences is the impairment of the ubiquitin/proteasome pathway, which leads to an accumulation of ubiquitinated proteins. If these proteins cannot be cleared, the cells activate apoptosis. Caspase-mediated proteolysis results in the cleavage of tau (and other proteins), and this generates aggregation-prone protein fragments that, if not cleared by the autophagy/lysosomal pathway, aggregate and promote neurodegeneration. Injured neurons, in turn, initiate the production of signals (eg, extracellular adenosine triphosphate and β-amyloid plaques) that further propel glial activation. Over time, this feed-forward cycle of glial activation and neuronal injury results in progressive neurodegeneration leading to the development of symptomatic Alzheimer's disease.

Abbreviations

aa	Amino acid
Aβ	Amyloid beta peptide
AD	Alzheimer's disease
Akt	Protein kinase B
cdk5	Cyclin-dependent kinase 5
CHIP	Carboxyl terminus of heat shock protein 70–interacting protein
GSK	Glycogen synthase kinase
Hsp	Heat shock protein
JNK	c-Jun N-terminal kinase
MT	Microtubule
NFT	Neurofibrillary tangle
PHF	Paired helical filament
PIN-1	Protein interacting with NIMA (never in mitosis gene A) 1
PP	Protein phosphatase
UbB+1	Frameshift mutant of ubiquitin B

process. The elucidation of the mechanisms mediated by neurotoxic products of inflammation that lead to AD neurodegeneration will open up new and important possible targets for the pharmacological treatment of AD.[57]

LATE-BREAKING STUDIES

Two recent studies address the proteolytic events leading to tau pathology. One study[58] proposes that an impairment of the ubiquitin/proteasome pathway leading to the accumulation of ubiquitinated proteins is an early response to neuronal damage. If cells cannot clear the accumulated ubiquitinated proteins, apoptosis starts and triggers caspase activation, which leads to tau cleavage. A failure to remove these aggregation-prone tau fragments by cathepsins promotes tau pathology. The other study[59] suggests that chaperone-mediated autophagy is involved in the delivery of cleaved tau to lysosomes for additional cleavage. The impaired translocation of cleaved tau across the lysosomal membrane seems to promote tau oligomerization and tau aggregation at the membrane and thus interferes with lysosomal functioning. Both studies conclude that disruption of the lysosomal pathway plays a role in the final stages of tau pathology. However, as suggested by the first study, impairment of the ubiquitin/proteasome pathway, followed by caspase activation, seems to be an earlier event in the proteolytic cascade involved in tau pathology.

Impairment of the ubiquitin/proteasome pathway, followed by caspase activation, seems to be an earlier event in the proteolytic cascade involved in tau pathology.

CONCLUSION

AD is an age-related neurodegenerative disorder that is associated with neuroinflammation. Even though studies, models, and hypotheses have flourished over the last 2 decades, little is known about the beginning of the pathology, and when symptoms are detected, the neurodegeneration is so advanced that little can be done. NFTs are a pathological hallmark of AD. The major component of NFTs is tau, a microtubule-associated protein that is abundant in neurons and is highly soluble, yet in AD, it appears as abnormal aggregates. Tau is hyperphosphorylated, is truncated at Asp421, and aggregates into insoluble PHFs. The elucidation of the sequence of these and other tau posttranslational modifications and their contribution to neuronal cell death is highly significant for identifying the upstream steps that can be therapeutically targeted to prevent neurons from reaching a point of no return.

The dynamics of tau phosphorylation/dephosphorylation are a main focus of attention for scientists trying to find a way to regulate this process to avoid tau hyperphosphorylation and aggregation, which impairs its clearance. However, because tau is a target for so many kinases and phosphatases, understanding this process has proven to be very difficult. The challenge is to dissect which pathways render tau prone to hyperphosphorylation and aggregation, such as its cleavage by caspases or priming by a cascade of kinases. Resolving the pathways that lead to tau hyperphosphorylation and aggregation will provide a basis for designing more effective therapeutic strategies for AD.

Neuroinflammation with the ensuing gliosis could be an early manifestation of AD neurodegeneration. A better understanding of the mechanisms by which products of inflammation mediate neuronal injury may lead to more effective anti-inflammatory therapeutic strategies for preventing or treating AD neurodegeneration, which is associated with chronic inflammation.

ACKNOWLEDGMENT

This work was supported by the National Institutes of Health [NINDS-NS41073 through the Specialized Neuroscience Research Program (Maria E. Figueiredo-Pereira is the head of the subproject) and NCRR-RR03037 (infrastructure) to Hunter College, City University of New York].

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Gotz J, Schild A, Hoerndli F, Pennanen L. Amyloid-induced neurofibrillary tangle formation in Alzheimer's disease: insight from transgenic mouse and tissue-culture models. *Int J Dev Neurosci* 2004; 22: 453–465.
2. Mandelkow E, von Bergen M, Biernat J, Mandelkow EM. Structural principles of tau and the paired helical filaments of Alzheimer's disease. *Brain Pathol* 2007; 17: 83–90.
3. Rosenberg KJ, Ross JL, Feinstein HE, et al. Complementary dimerization of microtubule-associated tau protein: implications for microtubule bundling and tau-mediated pathogenesis. *Proc Natl Acad Sci U S A* 2008; 105: 7445–7450.
4. Goedert M, Spillantini MG, Jakes R, et al. Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron* 1989; 3: 519–526.
5. Crowther T, Goedert M, Wischik CM. The repeat region of microtubule-associated protein tau forms part of the core of the paired helical filament of Alzheimer's disease. *Ann Med* 1989; 21: 127–132.
6. Spillantini MG, Goedert M. Tau protein pathology in neurodegenerative diseases. *Trends Neurosci* 1998; 21: 428–433.
7. Corsetti V, Amadoro G, Gentile A, et al. Identification of a caspase-derived N-terminal tau fragment in cellular and animal Alzheimer's disease models. *Mol Cell Neurosci* 2008; 38: 381–392.
8. Gomez-Ramos A, Smith MA, Perry G, Avila J. Tau phosphorylation and assembly. *Acta Neurobiol Exp (Wars)* 2004; 64: 33–39.
9. Johnson GV, Stoothoff WH. Tau phosphorylation in neuronal cell function and dysfunction. *J Cell Sci* 2004; 117: 5721–5729.
10. Tian Q, Wang J. Role of serine/threonine protein phosphatase in Alzheimer's disease. *Neurosignals* 2002; 11: 262–269.
11. Sengupta A, Kabat J, Novak M, et al. Phosphorylation of tau at both Thr 231 and Ser 262 is required for maximal inhibition of its binding to microtubules. *Arch Biochem Biophys* 1998; 357: 299–309.
12. Wang JZ, Grundke-Iqbal I, Iqbal K. Kinases and phosphatases and tau sites involved in Alzheimer neurofibrillary degeneration. *Eur J Neurosci* 2007; 25: 59–68.
13. Lu KP, Zhou XZ. The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nat Rev Mol Cell Biol* 2007; 8: 904–916.
14. Lu PJ, Wulf G, Zhou XZ, et al. The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* 1999; 399: 784–788.
15. Lu KP, Liou YC, Vincent I. Proline-directed phosphorylation and isomerization in mitotic regulation and in Alzheimer's disease. *Bioessays* 2003; 25: 174–181.
16. Illenberger S, Zheng-Fischhofer Q, Preuss U, et al. The endogenous and cell cycle-dependent phosphorylation of tau protein in living cells: implications for Alzheimer's disease. *Mol Biol Cell* 1998; 9: 1495–1512.
17. Billingsley ML, Kincaid RL. Regulated phosphorylation and dephosphorylation of tau protein: effects on microtubule interaction, intracellular trafficking and neurodegeneration. *Biochem J* 1997; 323(pt 3): 577–591.
18. Kosik KS, Qiu WQ, Greenberg S. Cellular signaling pathways and cytoskeletal organization. *Ann N Y Acad Sci* 1996; 777: 114–120.
19. Maudsley S, Mattson MP. Protein twists and turns in Alzheimer disease. *Nat Med* 2006; 12: 392–393.
20. Gamblin TC, Chen F, Zambrano A, et al. Caspase cleavage of tau: linking amyloid and neurofibrillary tangles in Alzheimer's disease. *Proc Natl Acad Sci U S A* 2003; 100: 10032–10037.
21. Eckert A, Keil U, Marques CA, et al. Mitochondrial dysfunction, apoptotic cell death, and Alzheimer's disease. *Biochem Pharmacol* 2003; 66: 1627–1634.
22. Rissman RA, Poon WW, Blurton-Jones M, et al. Caspase-cleavage of tau is an early event in Alzheimer

disease tangle pathology. *J Clin Invest* 2004; 114: 121–130.

23. Guillozet-Bongaarts AL, Cahill ME, Cryns VL, et al. Pseudophosphorylation of tau at serine 422 inhibits caspase cleavage: in vitro evidence and implications for tangle formation in vivo. *J Neurochem* 2006; 97: 1005–1014.
24. Rohn TT, Rissman RA, Davis MC, et al. Caspase-9 activation and caspase cleavage of tau in the Alzheimer's disease brain. *Neurobiol Dis* 2002; 11: 341–354.
25. Berry RW, Abraha A, Lagalwar S, et al. Inhibition of tau polymerization by its carboxy-terminal caspase cleavage fragment. *Biochemistry* 2003; 42: 8325–8331.
26. Dickson DW. Apoptotic mechanisms in Alzheimer neurofibrillary degeneration: cause or effect? *J Clin Invest* 2004; 114: 23–27.
27. Mondragon-Rodriguez S, Mena R, Binder LI, et al. Conformational changes and cleavage of tau in Pick bodies parallel the early processing of tau found in Alzheimer pathology. *Neuropathol Appl Neurobiol* 2008; 34: 62–75.
28. Wray S, Saxton M, Anderton BH, Hanger DP. Direct analysis of tau from PSP brain identifies new phosphorylation sites and a major fragment of N-terminally cleaved tau containing four microtubule-binding repeats. *J Neurochem.* doi:10.1111/j.1471–4159.2008.05321.x. Available at: http://www3.interscience.wiley.com.
29. Rapoport M, Dawson HN, Binder LI, et al. Tau is essential to beta-amyloid-induced neurotoxicity. *Proc Natl Acad Sci U S A* 2002; 99: 6364–6369.
30. Borsello T, Forloni G. JNK signalling: a possible target to prevent neurodegeneration. *Curr Pharm Des* 2007; 13: 1875–1886.
31. Song JJ, Lee YJ. Differential activation of the JNK signal pathway by UV irradiation and glucose deprivation. *Cell Signal* 2007; 19: 563–572.
32. Sahara N, Murayama M, Lee B, et al. Active c-jun N-terminal kinase induces caspase cleavage of tau and additional phosphorylation by GSK-3beta is required for tau aggregation. *Eur J Neurosci* 2008; 27: 2897–2906.
33. Cho SG, Choi EJ. Apoptotic signaling pathways: caspases and stress-activated protein kinases. *J Biochem Mol Biol* 2002; 35: 24–27.
34. Pei JJ, Braak E, Braak H, et al. Localization of active forms of C-jun kinase (JNK) and p38 kinase in Alzheimer's disease brains at different stages of neurofibrillary degeneration. *J Alzheimers Dis* 2001; 3: 41–48.
35. Keller JN, Hanni KB, Markesbery WR. Impaired proteasome function in Alzheimer's disease. *J Neurochem* 2000; 75: 436–439.
36. van Leeuwen FW, de Kleijn DP, van den Hurk HH, et al. Frameshift mutants of beta amyloid precursor protein and ubiquitin-B in Alzheimer's and Down patients. *Science* 1998; 279: 242–247.
37. Lam YA, Pickart CM, Alban A, et al. Inhibition of the ubiquitin-proteasome system in Alzheimer's disease. *Proc Natl Acad Sci U S A* 2000; 97: 9902–9906.
38. Keck S, Nitsch R, Grune T, Ullrich O. Proteasome inhibition by paired helical filament-tau in brains of patients with Alzheimer's disease. *J Neurochem* 2003; 85: 115–122.
39. Dorval V, Fraser PE. Small ubiquitin-like modifier (SUMO) modification of natively unfolded proteins tau and alpha-synuclein. *J Biol Chem* 2006; 281: 9919–9924.
40. David DC, Layfield R, Serpell L, et al. Proteasomal degradation of tau protein. *J Neurochem* 2002; 83: 176–185.
41. Babu JR, Geetha T, Wooten MW. Sequestosome 1/p62 shuttles polyubiquitinated tau for proteasomal degradation. *J Neurochem* 2005; 94: 192–203.
42. Brown MR, Bondada V, Keller JN, et al. Proteasome or calpain inhibition does not alter cellular tau levels in neuroblastoma cells or primary neurons. *J Alzheimers Dis* 2005; 7: 15–24.
43. Shimura H, Schwartz D, Gygi SP, Kosik KS. CHIP-Hsc70 complex ubiquitinates phosphorylated tau and enhances cell survival. *J Biol Chem* 2004; 279: 4869–4876.
44. Johnson ES. Protein modification by SUMO. *Annu Rev Biochem* 2004; 73: 355–382.
45. Sahara N, Murayama M, Mizoroki T, et al. In vivo evidence of CHIP up-regulation attenuating tau aggregation. *J Neurochem* 2005; 94: 1254–1263.
46. Goryunov D, Liem RK. CHIP-ping away at tau. *J Clin Invest* 2007; 117: 590–592.
47. McClellan AJ, Tam S, Kaganovich D, Frydman J. Protein quality control: chaperones culling corrupt conformations. *Nat Cell Biol* 2005; 7: 736–741.
48. Dickey CA, Koren J, Zhang YJ, et al. Akt and CHIP coregulate tau degradation through coordinated interactions. *Proc Natl Acad Sci U S A* 2008; 105: 3622–3627.
49. Takashima A. GSK-3 is essential in the pathogenesis of Alzheimer's disease. *J Alzheimers Dis* 2006; 9: 309–317.
50. Iqbal K, Alonso AC, Grundke-Iqbal I. Cytosolic abnormally hyperphosphorylated tau but not paired helical filaments sequester normal MAPs and inhibit microtubule assembly. *J Alzheimers Dis* 2008; 14: 365–370.
51. Mandelkow EM, Thies E, Trinczek B, et al. MARK/PAR1 kinase is a regulator of microtubule-dependent transport in axons. *J Cell Biol* 2004; 167: 99–110.
52. Dickey CA, Kamal A, Lundgren K, et al. The high-affinity HSP90-CHIP complex recognizes and selectively degrades phosphorylated tau client proteins. *J Clin Invest* 2007; 117: 648–658.
53. McGeer PL, McGeer EG. NSAIDs and Alzheimer disease: epidemiological, animal model and clinical studies. *Neurobiol Aging* 2007; 28: 639–647.
54. Sastre M, Walter J, Gentleman SM. Interactions between APP secretases and inflammatory mediators. *J Neuroinflamm* 2008; 5: 25.
55. Yoshiyama Y, Higuchi M, Zhang B, et al. Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model. *Neuron* 2007; 53: 337–351.
56. Arnaud L, Robakis NK, Figueiredo-Pereira ME. It may take inflammation, phosphorylation and ubiquitination to 'tangle' in Alzheimer's disease. *Neurodegener Dis* 2006; 3: 313–319.
57. Klegeris A, McGeer EG, McGeer PL. Therapeutic approaches to inflammation in neurodegenerative disease. *Curr Opin Neurol* 2007; 20: 351–357.
58. Arnaud LT, Myeku N, Figueiredo-Pereira ME. Proteasome-caspase-cathepsin sequence leading to tau pathology induced by prostaglandin J2 in neuronal cells. *J Neurochem* 2009; 110: 328–342.
59. Wang Y, Martinez-Vicente M, Kruger U, et al. Tau fragmentation, aggregation and clearance: the dual role of lysosomal processing. *Hum Mol Genet* 2009; 18: 4153–4170.

Cell "Self-Eating" (Autophagy) Mechanism in Alzheimer's Disease

Sarah F. Funderburk, PhD, Bridget K. Marcellino, BA, and Zhenyu Yue, PhD

Department of Neurology, Mount Sinai School of Medicine, New York, NY
Department of Neuroscience, Mount Sinai School of Medicine, New York, NY

OUTLINE

Beginnings of an Alzheimer's Disease–Autophagy Connection: Endosomal-Lysosomal Pathway
Amyloid-β Amyloidogenesis in Autophagosomes
Aberrant Induction and a Dysfunctional Autophagic Pathway
  Presenilin 1 Studies
  Oxidative Stress
  Okadaic Acid Model
Blocking Autophagy Degradation
  Impaired Clearance
  Tau and Autophagic Dysfunction
Role for Reduction of Autophagy
  Beclin 1–Deficient Mouse Model
  Herpes Simplex Virus Type 1
Late breaking studies and future perspectives

ABSTRACT

The autophagy pathway is the major degradation pathway of the cell for long-lived proteins and organelles. Dysfunction of autophagy has been linked to several neurodegenerative disorders that are associated with an accumulation of misfolded protein aggregates. Alzheimer's disease, the most common neurodegenerative disorder, is characterized by 2 aggregate forms, tau tangles and amyloid-β plaques. Autophagy has been linked to Alzheimer's disease pathogenesis through its merger with the endosomal-lysosomal system, which has been shown to play a role in the formation of the latter amyloid-β plaques. However, the precise role of autophagy in Alzheimer's disease pathogenesis is still under contention. One hypothesis is that aberrant autophagy induction results in an accumulation of autophagic vacuoles containing amyloid-β and the components necessary for its generation, whereas other evidence points to impaired autophagic clearance or even an overall reduction in autophagic activity playing a role in Alzheimer's disease pathogenesis. In this review, we discuss the current evidence linking autophagy to Alzheimer's disease as well as the uncertainty over the exact role and level of autophagic regulation in the pathogenic mechanism of Alzheimer's disease. ***Mt Sinai J Med 77:59–68, 2010.*** *© 2010 Mount Sinai School of Medicine*

Address Correspondence to:

Zhenyu Yue
Department of Neurology
Mount Sinai School of Medicine
New York, NY
Email: zhenyu.yue@mssm.edu

Key Words: Alzheimer's disease, amyloidogenesis, autophagy, dysfunction, endosomal-lysosomal pathway.

The cellular process of autophagy, called the waste management or housekeeping system of the cell,[1] eliminates cytoplasmic "waste" material and aids in the recycling of some of its components. The autophagic pathway is responsible for the delivery of cellular debris such as damaged organelles and obsolete intracellular macromolecules to lysosomes for degradation.[2,3] The word *autophagy* is derived from the Greek words *auto* (meaning "self") and *phage* (meaning "eat"), which precisely define the intent of autophagy to clear cellular contents.[4] There are 3 subtypes of autophagy distinguishable by their mechanisms and cellular functions: macroautophagy, microautophagy, and chaperone-mediated autophagy.[5–8]

The most prevalent form of autophagy is macroautophagy, which is discussed in this review and hereafter will be referred to as autophagy. Autophagy is the main mechanism used in eukaryotes for the catabolism of organelles and long-lived proteins. Such cytoplasmic debris is brought to lysosomes through double-membraned vacuoles known as autophagosomes. The cellular process of autophagy involves a series of ordered steps (Figure 1). Morphologically, the first step is the

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20161

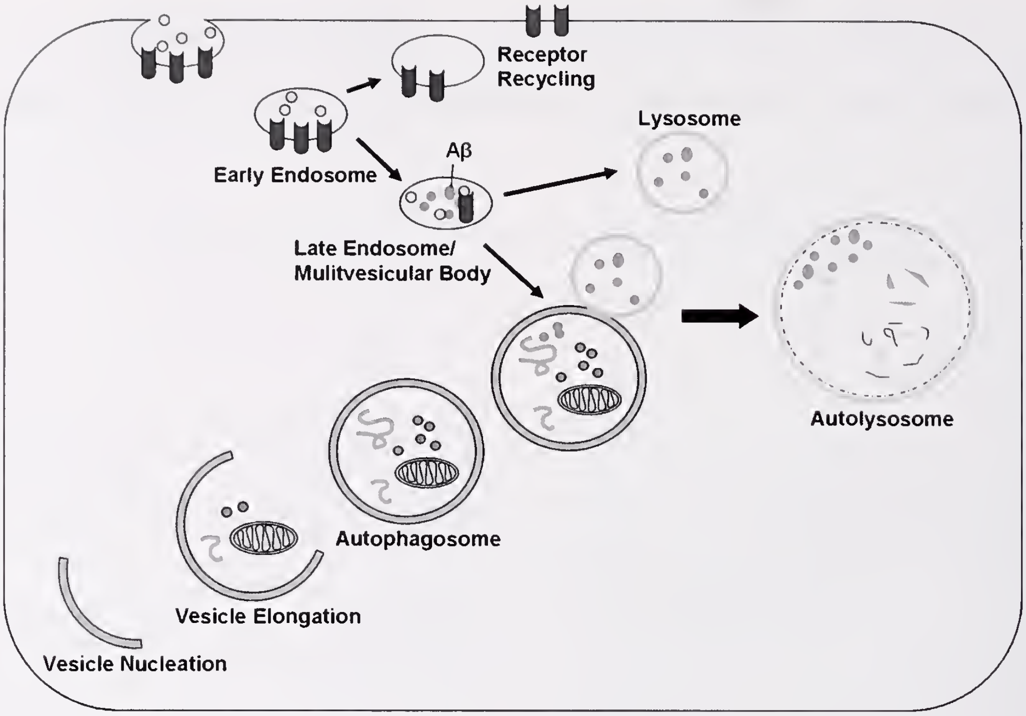


Fig 1. The endosomal-lysosomal pathway merges with the autophagic pathway. The autophagic pathway is used for organelle and protein turnover. Upon the induction of autophagy, a double-membraned structure, which is called the *isolation membrane*, forms by vesicle nucleation around cytoplasmic contents and becomes an autophagosome. Cytoplasmic substrates are then degraded after fusion of the autophagosome with the lysosome, and this results in the autolysosome. The endosomal pathway, which is important for the uptake and recycling of nutrients and transmembrane receptors, merges with the autophagic pathway. Aβ is generated in multivesicular bodies of the endosomal pathway and may also be generated in autophagosomes. **Abbreviation:** Aβ, amyloid-β.

nucleation process, by which an isolation membrane, also known as a phagophore, is formed. The membrane then elongates, and the edges of the membrane fuse; this results in the formation of enclosed vacuoles termed autophagosomes, which hold cytoplasmic contents. Lastly, the autophagosome fuses with a lysosome, which forms an autolysosome in which the engulfed material and the inner membrane are degraded.[8–10]

Autophagy has roles in both normal cellular homeostasis and disease states. The most studied form of autophagy is its induced response to a variety of stress conditions including nutrient starvation. The activation of autophagy results in enhanced protein degradation accompanied by an increase in the amino acid pool, which provides an energy source and allows for necessary protein synthesis.[11] In addition to its response to nutrient availability, a growing body of evidence indicates that autophagy is involved in an array of normal physiological processes as well as pathological conditions, such as development, cell death, aging, antigen presentation, bacterial degradation, and tumor suppression.[7,12] Most, if not all, mammalian cells also maintain a basal level of autophagy that is responsible for the normal intracellular turnover of proteins and organelles. This constitutively active autophagy (basal level) is believed to be especially critical for cells that are postmitotic, such as neurons, possibly because of their inability to dilute aberrant components through cell division.[13–15]

Despite the rapid growth of autophagy study, the knowledge of the autophagic process in neurons is extremely limited. Recent studies, demonstrating that loss of basal autophagy in the central nervous system causes an accumulation of ubiquitinated protein inclusions and neurodegeneration,[13,15] suggest an essential role for autophagy in the continuous removal of proteins destined for degradation and neuroprotection. Emerging evidence has indeed implicated autophagy in several major neurodegenerative disorders, such as frontotemporal dementia, Alzheimer's disease (AD), Parkinson's disease, and Huntington's disease.[16–19] Although many studies have suggested a neuroprotective role, the exact way in which autophagy is connected to these

Recent studies, demonstrating that loss of basal autophagy in the central nervous system causes an accumulation of ubiquitinated protein inclusions and neurodegeneration, suggest an essential role for autophagy in the continuous removal of proteins destined for degradation and neuroprotection.

neuropathological conditions is unclear. Consistent with a beneficial, cytoprotective function of autophagy, studies have shown that inhibition of autophagy results in an accumulation of α-synuclein in Parkinson's disease[20–22] and increased formation of huntingtin aggregates in Huntington's disease,[23] whereas increases in autophagic activity can aid in the clearance of these toxic aggregate species.[24] The involvement of autophagy in the clearance of protein aggregates associated with neurodegenerative disease is of great interest and has stimulated studies that seek the potential therapeutic value of autophagy in the treatment of neurodegenerative diseases associated with protein aggregates. However, some reports, questioning the universality of neuroprotective autophagy, have suggested that inappropriate autophagic induction contributes to cytotoxicity in AD.[25]

Of particular interest in the link of autophagy to disease pathogenesis is a series of studies that suggest dysfunctional autophagy in AD.[25–28] We review the current evidence that links autophagy to AD. In addition, we also discuss the uncertainty over the precise role that autophagy may play in the context of AD pathogenesis.

Of particular interest in the link of autophagy to disease pathogenesis is a series of studies that suggest dysfunctional autophagy in Alzheimer's disease (AD).

BEGINNINGS OF AN ALZHEIMER'S DISEASE–AUTOPHAGY CONNECTION: ENDOSOMAL-LYSOSOMAL PATHWAY

AD is the most common neurodegenerative disorder;[29] its symptomatic hallmarks include slow progression of memory impairment and loss of cognitive function. The AD-affected brain is characterized by 2 forms of aggregates: the intracellular tau tangle and the extracellular amyloid-β (Aβ) peptide neuritic plaque. The former is composed of tau, a microtubule-associated protein that is aberrantly phosphorylated in AD and aggregates into paired helical filaments,[30] whereas the latter is derived from proteolysis of amyloid precursor protein (APP).[31] Mutations in APP[32] and presenilin,[33,34] a protein involved in APP-to-Aβ proteolysis, cause familial Alzheimer's disease (FAD), which is rare and autosomal-dominant. Sporadic AD, which is more prevalent, shows many of the pathological characteristics of FAD, and this suggests that factors affecting the APP-to-Aβ pathway and subsequent Aβ-amyloidogenesis play a significant role in sporadic forms of the disease.[35]

The investigation of the role of autophagy in AD began because of the initial findings of the involvement of the endosomal-lysosomal system in the pathogenesis of AD and specifically Aβ-amyloidogenesis. Several studies have recognized the endosomal-lysosomal pathway as an important regulator of the processing of APP.[36–38] The endocytic pathway enables a cell to sample its surrounding environment and plays a critical role in the uptake of extracellular nutrients and regulation of cell-surface receptor expression. Sorting of internalized molecules occurs in the early endosome, which directs the material back to the plasma membrane for recycling, to the trans-Golgi network for further processing, or to late endosomes and lysosomes for degradation.[39] APP potentially undergoes processing at each of these locations. Early endosomes produce Aβ from APP in normal cells and mediate the uptake of Aβ and soluble APP. In the AD brain, activation of the endocytic pathway is the earliest noted intracellular manifestation of the disease, and neurons in susceptible brain areas exhibit progressive abnormalities in the endocytic pathway that include increases in the size and volume of early endosomes.[40] Aβ has also been detected in these enlarged endosomes that are immunopositive for the early endosomal marker rab5.[41]

Abnormalities in the lysosomal pathway also occur early in AD pathogenesis before the appearance of neurofibrillary tangles or senile plaques. An up-regulation in the lysosmal system occurs in vulnerable cell populations and results in increased numbers of lysosomes with elevated expression of lysosomal hydrolases.[42] These hydrolases include cathepsins, which have been shown to be both directly[43] and indirectly[39,44] involved in Aβ formation. As AD pathogenesis progresses, lysosomal dysfunction appears to occur with the buildup of vacuolar structures and the accumulation of $A\beta_{1-42}$. The degeneration of the compromised neurons leads to the release of these structures into the extracellular space, where they associate with deposits of Aβ.[45]

Although endocytic pathway induction and dysfunction can lead to lysosomal abnormalities, the autophagic pathway also converges on the lysosome (Figure 1). Additionally, the endocytic pathway in neurons exhibits a significant degree of merger with the autophagic pathway with the fusion of late endosomes with autophagosomes.[46] However, only in the last 10 years with advances in knowledge of the autophagic process has AD research begun to focus on the pathological effects of possible irregularities in the autophagic pathway.[39,47]

AMYLOID-β AMYLOIDOGENESIS IN AUTOPHAGOSOMES

A long-recognized characteristic feature of the AD brain is the presence of granulovacuolar structures in dystrophic (swollen) neurites.[48] As previously mentioned, the presence of these structures has been associated with an up-regulation/dysfunction of the endosomal-lysosomal pathway, as these structures tend to be lysosome-dense bodies that are acid phosphatase–positive and abundant in cathepsins, lysosomal proteases. More recent studies have identified many of these accumulated bodies as autophagosomes and autophagolysosomes, or early and late autophagic vacuoles (AVs), both intermediate structures of autophagy.[28] What could be of particular importance is the claim that these numerous AVs contain the components for Aβ generation, APP, and its processing enzymes.[25,49] APP has 3 sites of proteolytic cleavage (α, β, and γ), and Aβ is generated from a β-cleaved C-terminal APP fragment that subsequently undergoes γ-cleavage to create 40- to 42-mer Aβ peptides. The γ-secretase protein complex that includes presenilin mediates this γ-cleavage and has been shown to be highly active in accumulated AVs.

ABERRANT INDUCTION AND A DYSFUNCTIONAL AUTOPHAGIC PATHWAY

Presenilin 1 Studies

One theory linking autophagy and AD is that irregular autophagy stimulation results in increased Aβ production through the accumulation of AVs containing the γ-secretase machinery.[25] As a component of the γ-secretase complex, the transmembrane protein presenilin 1 (PS1) plays a significant role in the production of Aβ.[50] Specifically, PS1 has been found to regulate the processing of APP C-terminal fragments in the endoplasmic reticulum and Golgi.[51] In FAD, the most common mutations are found in PS1 and PS2. These mutations cause increased production of Aβ peptide leading to the accumulation of this protein and the formation of Aβ plaques.[52,53]

There is evidence suggesting a possible link between PS1 and AD through autophagy. The recent studies of PS1 mutated cell lines and mouse models have elucidated this potential link. In the PS1/APP mouse model of β-amyloidogenesis, there was an accumulation of AVs containing APP, PS1, and other proteins.[25] Similarly, in PS1-null neurons, there was a buildup of degradative vacuoles accompanied by both an abnormal presence of α-synuclein and β-synuclein proteins in these vacuoles and an increased level of both proteins. Interestingly, in control neurons, stimulation of autophagy caused a similar mislocalization of these synucleins. The effects seen in the PS1-deficient neurons were proposed to be due not to abnormal γ-secretase activity but possibly to abnormal regulation of calcium channels.[54]

There is evidence suggesting a possible link between transmembrane protein presenilin 1, which promotes the formation of Aβ plaques, and AD through autophagy.

Telencephalin (TLN), a neural intercellular adhesion molecule with which PS1 interacts,[55] has also been shown to accumulate in autophagic-like vacuoles in PS1-null hippocampal neurons. However, the prolonged half-life and enrichment of TLN were not a result of abnormal γ-secretase cleavage because TLN is not a substrate of γ-secretase. Rather, the authors hypothesized that the deficiency in PS1 caused a failure of autophagy to clear TLN by possibly disallowing proper vacuole and lysosome fusion.[56] All of the aforementioned evidence supports the proposition that impaired presenilin function translates into possible autophagic dysfunction either causing or exacerbating protein accumulation and, therefore, perhaps playing a significant role in the pathogenesis of neurodegenerative diseases such as AD.

Oxidative Stress

Oxidative damage to neurons has been reported to be an early occurrence in the pathogenesis of AD.[57] Zheng *et al.*[58,59] showed that the exposure of differentiated neuroblastoma cells to oxidative stress resulted in the accumulation of Aβ-enriched lysosomes. This Aβ accumulation was suppressed by 3-methyladenine (3-MA), an inhibitor of autophagosome formation, and this suggested the involvement of autophagy. The authors of the aforementioned study postulated that this increase in intralysosomal Aβ was due to oxidative stress inducing reparative autophagy, which resulted in the delivery of additional Aβ to autophagolysosomes. Other suggested mechanisms have included oxidative stress–induced enhancement of APP processing and decreased Aβ degradation due to either a decrease in lysosomal enzyme activity or a structural modification of Aβ that would make it more resistant to hydrolysis. Oxidative stress and damage in AD have also been proposed to lead to autophagy of mitochondria with an accumulation of mitochondrial degradation products also in the lysosomal pathway.[60–62]

Okadaic Acid Model

Okadaic acid (OA), a protein phosphatase-2A inhibitor, has been shown to be a useful tool for mimicking several pathological hallmarks of AD, including enhanced tau phosphorylation, Aβ deposition, and neuronal death. Use of OA is not sufficient for characterizing the mechanisms of the development of AD pathology, but it has been used to examine the role of tau phosphorylation in AD neuronal death.[63] A recent study demonstrated that treatment of rat cortical neurons with OA also resulted in the accumulation of AVs and elevated levels of the autophagosome marker microtubule-associated protein 1 light chain 3 II (LC3-II) as well as increases in the lysosomal system. OA both activated mammalian target of rapamycin, an inhibitor of autophagy, and up-regulated expression of the autophagy protein beclin 1, and this led the authors to conclude that the observed autophagic induction was due to the beclin 1–dependent autophagic pathway in their system. Blocking autophagy with 3-MA ameliorated OA-induced neurotoxicity.

BLOCKING AUTOPHAGY DEGRADATION

Impaired Clearance

Although several studies have implicated an induction of autophagy with the accumulation of AVs in AD models, all of these studies have also conceded that dysfunction in the autophagic pathway most likely plays a role. More specifically, the enrichment of AVs in AD suggests impairment of autophagic clearance.[27]

The role and level of autophagy in the normal brain are often disputed. Because of the relative dearth of AV intermediates in comparison to AD brain tissue, many contend that basal autophagic activity is low in the healthy brain. A recent study by Boland *et al.*,[26] however, has argued that autophagy in healthy neurons is extremely efficient, with autophagosomes being rapidly cleared by fusion with lysosomes. Whether basal activity is low or "efficiently" high, 2 separate studies have shown that autophagy is critical for neuronal survival. The deletion of 2 specific autophagy genes (autophagy-related 5 and autophagy-related 7) resulted in the blockage of constitutive autophagy and thereby promoted the formation of ubiquitinated protein aggregates, neurodegeneration, behavioral abnormalities, and early mortality.[13,15] The importance of a functional autophagy pathway was also affirmed by the aforementioned study by Boland and colleagues. In the study, rat cortical neurons were treated with rapamycin, lysosomal inhibitors, or vinblastine, which both induces autophagy and blocks AV clearance by lysosomes. Neurons treated with lysosomal inhibitors or vinblastine exhibited an accumulation of AVs with morphologies similar to those in the AD brain and the PS1/APP mouse model. The authors, therefore, concluded that autophagy dysfunction due to impaired AV clearance, and not autophagy induction alone, resulted in pathological patterns similar to AD.

> *Whether basal activity is low or "efficiently" high, 2 separate studies have shown that autophagy is critical for neuronal survival.*

> *The authors concluded that autophagy dysfunction due to impaired AV clearance, and not autophagy induction alone, resulted in pathological patterns similar to AD.*

Tau and Autophagic Dysfunction

Like Aβ, altered processing and impaired degradation by the lysosomal system have also been implicated in the buildup of tau neurofibrillary tangles in AD. At least 2 different tau transgenic mouse lines have exhibited autophagic-lysosomal abnormalities upon ultrastructural analysis.[64,65] The first study focused on transgenic mice expressing human tau with 4 tubulin-binding repeats [increased by frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17) splice donor mutations] and 3 FTDP-17 missense mutations: G272V, P301L, and R406W. An examination of these mice revealed an increase in lysosomal neuronal complexes and an up-regulation of acid phosphatases in regions that had accumulated filamentous tau aggregates.[65] In a second study, transgenic mice expressing mutant human (P301L) tau demonstrated axonal spheroids that were enriched with tau filaments as well as autophagosomes, with some spheroids appearing similar to those in the characteristic dystrophic neurites of AD.[64]

Overexpression of tau in other model systems has also demonstrated a possible link to autophagy dysfunction. In a study using an inducible tau expression system in cultured cells, the acidotropic agent chloroquine was used to disrupt the lysosomal system after tau induction, and this resulted in a delayed clearance of tau. Chloroquine treatment before and during tau induction led to an accumulation of tau with the formation of aggregates. This aggregation was not due to changes in tau phosphorylation, but an altered functioning of specific lysosomal proteases was found to be likely. Tau accumulation and aggregation were also induced when autophagy was inhibited in the cells with 3-MA treatment.[66] A second more recent study used an inducible mouse neuroblastoma N2a cell model featuring the repeat domain of tau with the FTDP-17 mutation and found that autophagy could degrade soluble mutant tau and its associated aggregates. Additionally, autophagy inhibition led to an increase in mutant tau aggregation and toxicity.[67] Similarly, induction of autophagy by rapamycin in *Drosophila* overexpressing wild-type and mutant tau resulted in increased tau clearance.[68] Like other aggregate-prone proteins, such as α-synuclein and huntingtin mutants, tau (or phosphorylated tau) can also be a substrate of autophagy; however, the mechanism underlying the degradation of specific disease-related proteins (likely in soluble form prior to the aggregates) is unclear.

Although never investigated, possible interference of dysfunctional tau with autophagosome clearance may contribute to AD neuropathology. Tau has been shown to regulate the stability of microtubules; therefore, modification of tau by phosphorylation or a change in the amount of tau can result in destabilization of neuronal microtubules. This can affect the placement and function of mitochondria, lysosomes, and other cellular structures and potentially put the cell in great danger.[69] Specifically, autophagosome clearance by lysosomes is a microtubule-dependent process.[70–72] Therefore, abnormal function of tau may disrupt neuronal function including axonal transport, which also involves trafficking of autophagic cargos,[73] and lead to the buildup of autophagosomes in AD dystrophic neurites or axons.[28]

ROLE FOR REDUCTION OF AUTOPHAGY

Beclin 1–Deficient Mouse Model

In what appears to be in opposition to some of the previously mentioned studies, recently there has also been evidence suggesting that decreased autophagic action may play a role in the pathogenic mechanisms of AD. Pickford *et al.*[74] reported that beclin 1 levels are decreased in the affected gray matter of patients with early stages of AD. Beclin 1 is part of a lipid kinase complex that is essential for autophagy and has been to shown to be involved in tumor suppression, neuroprotection, and cell death.[75–77] In contrast to the reduced cortical beclin 1 levels in AD patients, AD pathology in APP transgenic mice was not sufficient to cause a deficiency in beclin 1. However, APP transgenic mice heterozygous for beclin 1 demonstrated an increased accumulation of Aβ peptide, synaptic degeneration, and subsequent neuronal death. A reduction of autophagy in beclin 1–deficient mice was indicated by reduced levels of LC3-II in comparison with those of the control mice. Autophagic reduction in these mice was accompanied by an accumulation of disrupted lysosomal structures. A change in the microglial response to Aβ peptide suggested an increase in brain injury. Overexpression of beclin 1, perhaps by the stimulation of autophagy, resulted in a reduction of both extracellular and intracellular Aβ peptide.[74] In this set of experiments, reduction of autophagy appeared to potentiate AD pathology, and this suggests that autophagy may also be neuroprotective in AD. This would indicate autophagy stimulation as a potential therapeutic route for AD.

Herpes Simplex Virus Type 1

Herpes simplex virus type 1 (HSV1) was initially proposed as an AD causal agent because of the parallels between the brain regions affected by the virus in herpes simplex encephalitis and those affected in AD.[78] Nine years later, Jamieson *et al.*[79] reported that latent HSV1 DNA is present in a large proportion of AD patients and in normal elderly populations. Subsequent studies have further strengthened the idea of a possible connection by linking apolipoprotein E with the severity of HSV1 infection. Wozniak's group[80] has also claimed to have demonstrated that HSV1 infection can result in $A\beta$ deposition and phosphorylation of tau.

HSV1 can disrupt autophagy by counteracting the protein kinase R (PKR) defense mechanism, which is activated by the presence of double-stranded RNA. Activated PKR would normally enhance autophagy by phosphorylating elongation initiation factor 2α (eIF2α), and this would result in the termination of protein synthesis and the initiation of apoptosis. HSV1 interferes with the rescue action of PKR by dephosphorylating eIF2α.[81] Specifically, a protein encoded by HSV1, infected cell polypeptide 34.5, has been reported to bind beclin 1 and inhibit its autophagy function.[82] Autophagic interference by HSV1 has been proposed to prevent the degradation of $A\beta$ and neurofibrillary tangles that accumulate in AD,[81] and this is consistent with the hypothesis that impaired autophagy in neurons in part contributes to the neuropathology of AD.

LATE BREAKING STUDIES AND FUTURE PERSPECTIVES

Much recent work in the field of Alzheimer's disease has provided molecular and cellular evidence that links autophagy to the pathogenesis of AD.[26,63,66,74,81] These studies undoubtedly strengthen the mechanistic view that centralizes on an axis of endocytosis, autophagy, and lysosomal degradation in the cellular pathway of AD, especially given the emerging evidence showing the convergence of endosomes and autophagosomes in the lysosomal degradation route. Although these studies implicate aberrant autophagic activity in the process of $A\beta$ synthesis or deposits, they seem to suggest 2 competing hypotheses with respect to the levels of autophagosomes or autophagic activity in the pathological process of AD[27,74] (Figure 2).

Two studies that were published during the preparation of this review emphasize the polarized views of autophagy's role in $A\beta$-induced neurotoxicity.[83,84] Using *Drosophila* as a model system, Ling *et al.*[84] reported that flies expressing $A\beta_{1-42}$ demonstrated an age-dependent increase in autophagic dysfunction. Moreover, down-regulating autophagy in flies in which specific autophagy genes were either reduced in copy or silenced in

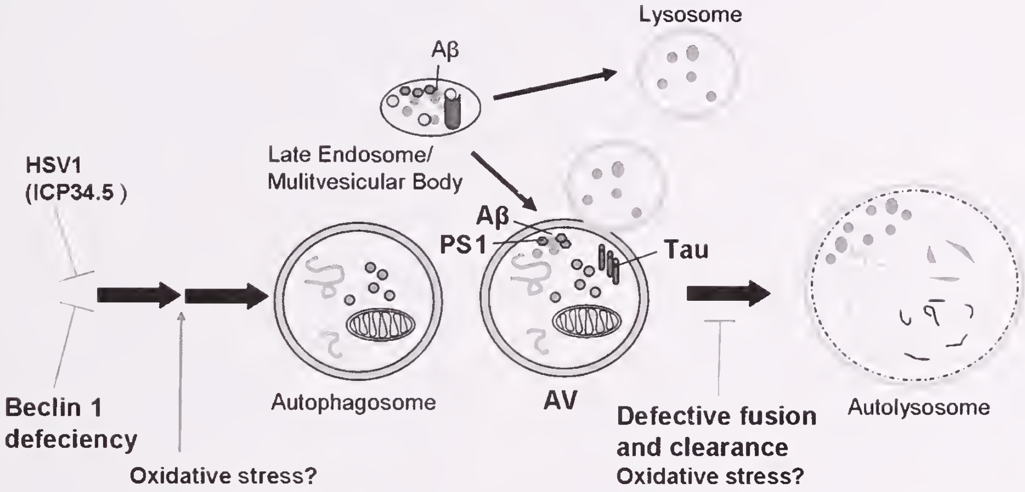


Fig 2. Possible alterations of the autophagic pathway lead to the accumulation of AVs in AD. Changes in the autophagic pathway (especially the accumulation of AVs in dystrophic neurites) have been linked to AD. The specific involvement of autophagy is unclear, but 2 competing hypotheses have emerged. One hypothesis links autophagic inhibition to AD pathogenesis. Autophagy could be inhibited in the AD brain by a reduction in beclin 1 or by the HSV1 protein ICP34.5, which binds and inhibits beclin 1. A second hypothesis focuses on aberrant autophagy induction and/or defective lysosomal fusion and clearance. Oxidative stress has been proposed to cause increases in $A\beta$ by up-regulating autophagy or by affecting its capacity for degradation and clearance. The mechanism for defective lysosomal fusion and subsequent clearance is not known, but the accumulation of intracellular $A\beta$ or tau may play a role. **Abbreviations:** $A\beta$, amyloid-β; AD, Alzheimer's disease; AV, autophagic vacuole; HSV1, herpes simplex virus type 1; ICP34.5, infected cell polypeptide 34.5; PS1, presenilin 1.

selective neurons resulted in a protective effect on the $A\beta_{1-42}$ toxicity. In the opposite manner, $A\beta_{1-42}$ flies treated with rapamycin to induce autophagy experienced increased toxicity and a shortened lifespan. A second recent article, however, demonstrated protective effects of autophagy on $A\beta_{1-42}$ toxicity in a neuroblastoma cell model. Cells were treated with extracellular $A\beta$, which induced an autophagic process and resulted in the accumulation of reactive oxygen species. Cells given a cotreatment of rapamycin exhibited a decrease in reactive oxygen species accumulation and increased survival, whereas those treated with the autophagic inhibitor wortmannin did not display the same protective effects.[83]

Many factors, such as different animal models, cellular models, and experimental paradigms, can contribute to the differences in conclusions. Also, the different model systems are likely representative of different particular stages of the AD pathogenic process. Therefore, properly functioning autophagy as a protective mechanism in one study may not necessarily contradict the finding that enhanced dysfunctional autophagy is neurotoxic in a different study. However, the major impediment to the revelation of the exact mechanism is undoubtedly the lack of knowledge of the neuronal autophagic process. Future work will be needed not only to further investigate the details of altered autophagy in AD neurons but also to characterize basal autophagy function under normal conditions.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Levine B, Kroemer G. Autophagy in the pathogenesis of disease. *Cell* 2008; 132: 27–42.
2. Dunn WA Jr. Autophagy and related mechanisms of lysosome-mediated protein degradation. *Trends Cell Biol* 1994; 4: 139–143.
3. Mortimore GE, Poso AR, Lardeux BR. Mechanism and regulation of protein degradation in liver. *Diabetes Metab Rev* 1989; 5: 49–70.
4. de Duve D. The peroxisome: a new cytoplasmic organelle. *Proc R Soc Lond B Biol Sci* 1969; 173: 71–83.
5. Mizushima N, Levine B, Cuervo AM, Klionsky DJ. Autophagy fights disease through cellular self-digestion. *Nature* 2008; 451: 1069–1075.
6. Kaushik S, Cuervo AM. Chaperone-mediated autophagy. *Methods Mol Biol* 2008; 445: 227–244.
7. Mizushima N. The pleiotropic role of autophagy: from protein metabolism to bactericide. *Cell Death Differ* 2005; 12(suppl 2): 1535–1541.
8. Klionsky DJ, Ohsumi Y. Vacuolar import of proteins and organelles from the cytoplasm. *Annu Rev Cell Dev Biol* 1999; 15: 1–32.
9. Dunn WA Jr. Studies on the mechanisms of autophagy: maturation of the autophagic vacuole. *J Cell Biol* 1990; 110: 1935–1945.
10. Dunn WA Jr. Studies on the mechanisms of autophagy: formation of the autophagic vacuole. *J Cell Biol* 1990; 110: 1923–1933.
11. Mizushima N. Autophagy: process and function. *Genes Dev* 2007; 21: 2861–2873.
12. Levine B, Klionsky DJ. Development by self-digestion: molecular mechanisms and biological functions of autophagy. *Dev Cell* 2004; 6: 463–477.
13. Komatsu M, Waguri S, Chiba T, et al. Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature* 2006; 441: 880–884.
14. Komatsu M, Waguri S, Ueno T, et al. Impairment of starvation-induced and constitutive autophagy in Atg7-deficient mice. *J Cell Biol* 2005; 169: 425–434.
15. Hara T, Nakamura K, Matsui M, et al. Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* 2006; 441: 885–889.
16. Martinez-Vicente M, Cuervo AM. Autophagy and neurodegeneration: when the cleaning crew goes on strike. *Lancet Neurol* 2007; 6: 352–361.
17. Rusten TE, Simonsen A. ESCRT functions in autophagy and associated disease. *Cell Cycle* 2008; 7: 1166–1172.
18. Williams A, Jahreiss L, Sarkar S, et al. Aggregate-prone proteins are cleared from the cytosol by autophagy: therapeutic implications. *Curr Top Dev Biol* 2006; 76: 89–101.
19. Rubinsztein DC, Gestwicki JE, Murphy LO, Klionsky DJ. Potential therapeutic applications of autophagy. *Nat Rev Drug Discov* 2007; 6: 304–312.
20. Vogiatzi T, Xilouri M, Vekrellis K, Stefanis L. Wild type alpha-synuclein is degraded by chaperone-mediated autophagy and macroautophagy in neuronal cells. *J Biol Chem* 2008; 283: 23542–23556.
21. Webb JL, Ravikumar B, Atkins J, et al. Alpha-synuclein is degraded by both autophagy and the proteasome. *J Biol Chem* 2003; 278: 25009–25013.
22. Paxinou E, Chen Q, Weisse M, et al. Induction of alpha-synuclein aggregation by intracellular nitrative insult. *J Neurosci* 2001; 21: 8053–8061.
23. Ravikumar B, Duden R, Rubinsztein DC. Aggregate-prone proteins with polyglutamine and polyalanine expansions are degraded by autophagy. *Hum Mol Genet* 2002; 11: 1107–1117.
24. Yamamoto A, Cremona ML, Rothman JE. Autophagy-mediated clearance of huntingtin aggregates triggered by the insulin-signaling pathway. *J Cell Biol* 2006; 172: 719–731.
25. Yu WH, Cuervo AM, Kumar A, et al. Macroautophagy–a novel beta-amyloid peptide-generating pathway activated in Alzheimer's disease. *J Cell Biol* 2005; 171: 87–98.
26. Boland B, Kumar A, Lee S, et al. Autophagy induction and autophagosome clearance in neurons: relationship to autophagic pathology in Alzheimer's disease. *J Neurosci* 2008; 28: 6926–6937.
27. Nixon RA. Autophagy, amyloidogenesis and Alzheimer disease. *J Cell Sci* 2007; 120(pt 23): 4081–4091.

28. Nixon RA, Wegiel J, Kumar A, et al. Extensive involvement of autophagy in Alzheimer disease: an immuno-electron microscopy study. *J Neuropathol Exp Neurol* 2005; 64: 113–122.
29. Ferri CP, Prince M, Brayne C, et al. Global prevalence of dementia: a Delphi consensus study. *Lancet* 2005; 366: 2112–2117.
30. Goedert M. Tau protein and the neurofibrillary pathology of Alzheimer's disease. *Trends Neurosci* 1993; 16: 460–465.
31. Glenner GG, Wong CW. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem Biophys Res Commun* 1984; 120: 885–890.
32. Goate A, Chartier-Harlin MC, Mullan M, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 1991; 349: 704–706.
33. Levy-Lahad E, Wasco W, Poorkaj P, et al. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* 1995; 269: 973–977.
34. Sherrington R, Rogaev EI, Liang Y, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 1995; 375: 754–760.
35. Hardy J. Amyloid, the presenilins and Alzheimer's disease. *Trends Neurosci* 1997; 20: 154–159.
36. Grbovic OM, Mathews PM, Jiang Y, et al. Rab5-stimulated up-regulation of the endocytic pathway increases intracellular beta-cleaved amyloid precursor protein carboxyl-terminal fragment levels and Abeta production. *J Biol Chem* 2003; 278: 31261–31268.
37. Koo EH, Squazzo SL. Evidence that production and release of amyloid beta-protein involves the endocytic pathway. *J Biol Chem* 1994; 269: 17386–17389.
38. Pasternak SH, Bagshaw RD, Guiral M, et al. Presenilin-1, nicastrin, amyloid precursor protein, and gamma-secretase activity are co-localized in the lysosomal membrane. *J Biol Chem* 2003; 278: 26687–26694.
39. Nixon RA, Cataldo AM, Mathews PM. The endosomal-lysosomal system of neurons in Alzheimer's disease pathogenesis: a review. *Neurochem Res* 2000; 25: 1161–1172.
40. Cataldo AM, Barnett JL, Pieroni C, Nixon RA. Increased neuronal endocytosis and protease delivery to early endosomes in sporadic Alzheimer's disease: neuropathologic evidence for a mechanism of increased beta-amyloidogenesis. *J Neurosci* 1997; 17: 6142–6151.
41. Cataldo AM, Petanceska S, Terio NB, et al. Abeta localization in abnormal endosomes: association with earliest Abeta elevations in AD and Down syndrome. *Neurobiol Aging* 2004; 25: 1263–1272.
42. Cataldo AM, Barnett JL, Berman SA, et al. Gene expression and cellular content of cathepsin D in Alzheimer's disease brain: evidence for early up-regulation of the endosomal-lysosomal system. *Neuron* 1995; 14: 671–680.
43. Chevallier N, Vizzavona J, Marambaud P, et al. Cathepsin D displays in vitro beta-secretase-like specificity. *Brain Res* 1997; 750: 11–19.
44. Beyreuther K, Pollwein P, Multhaup G, et al. Regulation and expression of the Alzheimer's beta/A4 amyloid protein precursor in health, disease, and Down's syndrome. *Ann N Y Acad Sci* 1993; 695: 91–102.
45. Cataldo AM, Hamilton DJ, Nixon RA. Lysosomal abnormalities in degenerating neurons link neuronal compromise to senile plaque development in Alzheimer disease. *Brain Res* 1994; 640: 68–80.
46. Larsen KE, Sulzer D. Autophagy in neurons: a review. *Histol Histopathol* 2002; 17: 897–908.
47. Mizushima N, Ohsumi Y, Yoshimori T. Autophagosome formation in mammalian cells. *Cell Struct Funct* 2002; 27: 421–429.
48. Suzuki K, Terry RD. Fine structural localization of acid phosphatase in senile plaques in Alzheimer's presenile dementia. *Acta Neuropathol* 1967; 8: 276–284.
49. Yu WH, Kumar A, Peterhoff C, et al. Autophagic vacuoles are enriched in amyloid precursor protein-secretase activities: implications for beta-amyloid peptide over-production and localization in Alzheimer's disease. *Int J Biochem Cell Biol* 2004; 36: 2531–2540.
50. Xia W, Ostaszewski BL, Kimberly WT, et al. FAD mutations in presenilin-1 or amyloid precursor protein decrease the efficacy of a gamma-secretase inhibitor: evidence for direct involvement of PS1 in the gamma-secretase cleavage complex. *Neurobiol Dis* 2000; 7(pt B): 673–681.
51. Xia W, Zhang J, Ostaszewski BL, et al. Presenilin 1 regulates the processing of beta-amyloid precursor protein C-terminal fragments and the generation of amyloid beta-protein in endoplasmic reticulum and Golgi. *Biochemistry* 1998; 37: 16465–16471.
52. Xia W. Role of presenilin in gamma-secretase cleavage of amyloid precursor protein. *Exp Gerontol* 2000; 35: 453–460.
53. Gandy S. The role of cerebral amyloid beta accumulation in common forms of Alzheimer disease. *J Clin Invest* 2005; 115: 1121–1129.
54. Wilson CA, Murphy DD, Giasson BI, et al. Degradative organelles containing mislocalized alpha-and beta-synuclein proliferate in presenilin-1 null neurons. *J Cell Biol* 2004; 165: 335–346.
55. Gahmberg CG. Leukocyte adhesion: CD11/CD18 integrins and intercellular adhesion molecules. *Curr Opin Cell Biol* 1997; 9: 643–650.
56. Esselens C, Oorschot V, Baert V, et al. Presenilin 1 mediates the turnover of telencephalin in hippocampal neurons via an autophagic degradative pathway. *J Cell Biol* 2004; 166: 1041–1054.
57. Zhu X, Raina AK, Lee HG, et al. Oxidative stress signalling in Alzheimer's disease. *Brain Res* 2004; 1000: 32–39.
58. Zheng L, Roberg K, Jerhammar F, et al. Oxidative stress induces intralysosomal accumulation of Alzheimer amyloid beta-protein in cultured neuroblastoma cells. *Ann N Y Acad Sci* 2006; 1067: 248–251.
59. Zheng L, Marcusson J, Terman A. Oxidative stress and Alzheimer disease: the autophagy connection? *Autophagy* 2006; 2: 143–145.
60. Arendt T, Holzer M, Fruth R, et al. Phosphorylation of tau, Abeta-formation, and apoptosis after in vivo inhibition of PP-1 and PP-2A. *Neurobiol Aging* 1998; 19: 3–13.
61. Tanaka T, Zhong J, Iqbal K, et al. The regulation of phosphorylation of tau in SY5Y neuroblastoma cells: the role of protein phosphatases. *FEBS Lett* 1998; 426: 248–254.
62. Yoon S, Choi J, Yoon J, et al. Okadaic acid induces JNK activation, bim overexpression and mitochondrial

dysfunction in cultured rat cortical neurons. *Neurosci Lett* 2006; 394: 190–195.
63. Yoon SY, Choi JE, Kweon HS, et al. Okadaic acid increases autophagosomes in rat neurons: implications for Alzheimer's disease. *J Neurosci Res* 2008; 86: 3230–3239.
64. Lin WL, Lewis J, Yen SH, et al. Ultrastructural neuronal pathology in transgenic mice expressing mutant (P301L) human tau. *J Neurocytol* 2003; 32: 1091–1105.
65. Lim F, Hernandez F, Lucas JJ, et al. FTDP-17 mutations in tau transgenic mice provoke lysosomal abnormalities and tau filaments in forebrain. *Mol Cell Neurosci* 2001; 18: 702–714.
66. Hamano T, Gendron TF, Causevic E, et al. Autophagic-lysosomal perturbation enhances tau aggregation in transfectants with induced wild-type tau expression. *Eur J Neurosci* 2008; 27: 1119–1130.
67. Wang Y, Martinez-Vicente M, Kruger U, et al. Tau fragmentation, aggregation and clearance: the dual role of lysosomal processing. *Hum Mol Genet* 2009; 18: 4153–4170.
68. Berger Z, Ravikumar B, Menzies FM, et al. Rapamycin alleviates toxicity of different aggregate-prone proteins. *Hum Mol Genet* 2006; 15: 433–442.
69. Avila J, Lucas JJ, Perez M, Hernandez F. Role of tau protein in both physiological and pathological conditions. *Physiol Rev* 2004; 84: 361–384.
70. Matteoni R, Kreis TE. Translocation and clustering of endosomes and lysosomes depends on microtubules. *J Cell Biol* 1987; 105: 1253–1265.
71. Jahreiss L, Menzies FM, Rubinsztein DC. The itinerary of autophagosomes: from peripheral formation to kiss-and-run fusion with lysosomes. *Traffic* 2008; 9: 574–587.
72. Kovacs AL, Seglen PO. Inhibition of hepatocytic protein degradation by inducers of autophagosome accumulation. *Acta Biol Med Ger* 1982; 41: 125–130.
73. Yue Z. Regulation of neuronal autophagy in axon: implication of autophagy in axonal function and dysfunction/degeneration. *Autophagy* 2007; 3: 139–141.
74. Pickford F, Masliah E, Britschgi M, et al. The autophagy-related protein beclin 1 shows reduced expression in early Alzheimer disease and regulates amyloid beta accumulation in mice. *J Clin Invest* 2008; 118: 2190–2199.
75. Yue Z, Jin S, Yang C, et al. Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor. *Proc Natl Acad Sci U S A* 2003; 100: 15077–15082.
76. Cao Y, Klionsky DJ. Physiological functions of Atg6/beclin 1: a unique autophagy-related protein. *Cell Res* 2007; 17: 839–849.
77. Liang XH, Jackson S, Seaman M, et al. Induction of autophagy and inhibition of tumorigenesis by beclin 1. *Nature* 1999; 402: 672–676.
78. Ball MJ. Limbic predilection in Alzheimer dementia: is reactivated herpesvirus involved? *Can J Neurol Sci* 1982; 9: 303–306.
79. Jamieson GA, Maitland NJ, Wilcock GK, et al. Latent herpes simplex virus type 1 in normal and Alzheimer's disease brains. *J Med Virol* 1991; 33: 224–227.
80. Wozniak MA, Itzhaki RF, Shipley SJ, Dobson CB. Herpes simplex virus infection causes cellular beta-amyloid accumulation and secretase upregulation. *Neurosci Lett* 2007; 429: 95–100.
81. Itzhaki RF, Cosby SL, Wozniak MA. Herpes simplex virus type 1 and Alzheimer's disease: the autophagy connection. *J Neurovirol* 2008; 14: 1–4.
82. Orvedahl A, Alexander D, Talloczy Z, et al. HSV-1 ICP34.5 confers neurovirulence by targeting the beclin 1 autophagy protein. *Cell Host Microbe* 2007; 1: 23–35.
83. Hung SY, Huang WP, Liou HC, Fu WM. Autophagy protects neuron from Abeta-induced cytotoxicity. *Autophagy* 2009; 5: 502–510.
84. Ling D, Song HJ, Garza D, et al. Abeta42-induced neurodegeneration via an age-dependent autophagic-lysosomal injury in *Drosophila*. *PLoS One* 2009; 4: e4201.

Transgenic Mouse Models of Alzheimer's Disease

Gregory A. Elder[1,3,4], Miguel A. Gama Sosa[2,3], and Rita De Gasperi[2,3]

[1]Neurology Service, James J. Peters Department of Veterans Affairs Medical Center, Bronx, NY
[2]Research and Development Service, James J. Peters Department of Veterans Affairs Medical Center, Bronx, NY
[3]Department of Psychiatry, Mount Sinai School of Medicine, New York, NY
[4]Department of Neurology, Mount Sinai School of Medicine, New York, NY

OUTLINE

- ALZHEIMER'S DISEASE AS A CANDIDATE DISEASE FOR MODELING IN TRANSGENIC ORGANISMS
- METHODS FOR GENERATING TRANSGENIC ORGANISMS
- TRANSGENIC MODELING BASED ON OVEREXPRESSION OF THE AMYLOID PRECURSOR PROTEIN
- TRANSGENIC MODELING WITH PRESENILIN MUTATIONS
- RELEVANCE AS MODELS OF ALZHEIMER'S DISEASE
- USE OF TRANSGENIC MODELS TO STUDY THE PATHOPHYSIOLOGY OF ALZHEIMER'S DISEASE: EFFECTS OF DIFFERENT β-AMYLOID SPECIES
- TRANSGENIC MODELS AND THE RISE OF THE β-AMYLOID OLIGOMER
- β-AMYLOID IMMUNIZATION AS A THERAPEUTIC STRATEGY IN ALZHEIMER'S DISEASE
- CONCLUSION

ABSTRACT

Alzheimer's disease is the most common cause of senile dementia in the United States and Europe. At present, there is no effective treatment. Given the disease's prevalence and poor prognosis, the development of animal models has been a high research priority. Transgenic modeling has been pursued on the basis of the amyloid hypothesis and has taken advantage of mutations in the amyloid precursor protein and the presenilins that cause familial forms of Alzheimer's disease. Modeling has been most aggressively pursued in mice, for which the techniques of genetic modification are well developed. Transgenic mouse models now exist that mimic a range of Alzheimer's disease–related pathologies. Although none of the models fully replicates the human disease, the models have contributed significant insights into the pathophysiology of β-amyloid toxicity, particularly with respect to the effects of different β-amyloid species and the possible pathogenic role of β-amyloid oligomers. They have also been widely used in the preclinical testing of potential therapeutic modalities and have played a pivotal role in the development of immunotherapies for Alzheimer's disease that are currently in clinical trials. These models will, without a doubt, continue to play central roles in preclinical testing and be used as tools for developing insights into the biological basis of Alzheimer's disease. ***Mt Sinai J Med 77:69–81, 2010.*** © *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Gregory A. Elder
Neurology Service
James J. Peters Department of
Veterans Affairs Medical Center
Bronx, NY
Email: gregory.elder@mssm.edu

Key Words: Alzheimer's disease, amyloid precursor protein, animal model, presenilins, transgenic mice.

Alzheimer's disease (AD), the most common cause of senile dementia in Europe and the United States, accounts for an estimated 50% to 80% of all cases. Clinically, AD is characterized by progressive cognitive impairment, usually beginning as memory loss but eventually progressing to involve multiple cognitive and behavioral domains. Pathologically, the disease is recognized by the presence of senile plaques, neurofibrillary tangles (NFTs), and neuronal loss. At present, there is no effective treatment. Given the prevalence and poor prognosis of the disease,

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20159

the development of animal models has been a high research priority.

ALZHEIMER'S DISEASE AS A CANDIDATE DISEASE FOR MODELING IN TRANSGENIC ORGANISMS

Transgenic technology exists for many organisms, including mice, rats, fish, flies, and worms. Modeling in invertebrates such as *Drosophila* and *Caenorhabditis elegans* offers advantages in terms of the high degree of experimental control and the relatively short life span of the organisms. They, however, suffer the disadvantage of being far removed phylogenetically from mammals. Although efforts to model AD in these systems continue, they have had less impact than models in mammalian systems and are not discussed further. The interested reader is, however, referred to 2 recent reviews of transgenic modeling of AD in *Drosophila* and *C. elegans*.[1,2] Among vertebrates, mice are by far the species most used. Transgenic modeling in mice is relatively inexpensive. Mice also have a relatively short life span, and the techniques for performing genetic modifications in them are well developed. Transgenic technology exists for rats but is not as widely available as that for mice, although some transgenic AD rat lines have recently been described.[3–6]

Regardless of the species chosen, transgenic technologies introduce genetic modifications. Therefore, successful modeling requires the disease to be associated with a genetic mutation or at least for a hypothesis to exist regarding the likely pathophysiology of the disorder that can be modeled by a genetic modification. To be useful as an animal model, the transgenic organism must also be able to exhibit the essential pathological, physiological, or behavioral features of the human disease.

AD may in many ways be regarded as the ideal disease for modeling in transgenic animals. First, it has a well-recognized pathology consisting of senile plaques and NFTs. The major constituents of these lesions are well defined, being the β-amyloid (Aβ) peptide in the case of plaques and hyperphosphorylated forms of tau in NFTs. AD also has other well-recognized pathological features, including neuronal and synaptic loss, dystrophic neurites, reactive astrocytes, and activated microglia. There is, in addition, a well-defined behavioral phenotype that can be modeled in the mouse.

> *Alzheimer's disease (AD) may in many ways be regarded as the ideal disease for modeling in transgenic animals.*

In AD, the 39– to 42–amino acid Aβ peptide deposits in senile plaques.[7] The Aβ peptide itself is derived from the processing of a larger precursor protein known as the amyloid precursor protein (APP). Many in vitro and in vivo studies have demonstrated that, in particular, the longer Aβ42 species can be neurotoxic. The amyloid cascade hypothesis postulates that shunting of amyloid precursor protein (APP) processing toward Aβ production sets off a chain of pathological events,[7] a process that can be modeled in transgenic animals. In addition, although most cases of AD occur sporadically, autosomal dominant forms of the disease exist that mimic the sporadic disease clinically and pathologically. Mutations in 3 genes including APP have been identified as causing familial Alzheimer's disease (FAD).[8] Indeed, the amyloid hypothesis was greatly bolstered by the finding that mutations in APP can cause FAD. These familial cases thus provide a genetic lesion that has been used to advantage in AD modeling in transgenic animals.

METHODS FOR GENERATING TRANSGENIC ORGANISMS

Transgenic organisms are generated by 1 of 2 general strategies. In the first, a genetic modification is introduced on top of the existing genetic makeup of the organism. In the second, the homologous gene of interest is modified selectively in its normal chromosomal position; this process is called gene targeting. These strategies have been developed to different degrees in different organisms.

In mice, both approaches are highly developed. In the first approach, a 1-cell embryo at the pronuclear stage is injected with a transgene. The transgene contains the coding region, often in the form of a complementary DNA of the protein of interest, coupled to a promoter that drives expression. The promoter is typically not the promoter of the native gene but rather a heterologous promoter chosen because of its strength or pattern of expression. However, transgenes may also consist of segments of genomic DNA that contain the gene of interest, often in the form of bacterial artificial chromosomes or P1 artificial chromosomes, in which case the transgene is driven by its native promoter and enhancers. The injected transgene integrates randomly, typically in multiple copies at a single site. Because no corresponding allele exists on the homologous chromosome opposite the integration site, these mice are usually called hemizygous. Because the heterologous promoters chosen are typically strong, the transgenic protein is often expressed at levels higher than would be present physiologically. In addition, because the mouse typically contains its own endogenous version of the gene, the transgene is generally expressed on top of the background expression of the mouse's endogenous gene, and this leads to further overexpression.

With gene targeting, instead of a foreign transgene being introduced, an endogenous gene in the mouse is modified. Initially, the modification is made in specialized cells termed embryonic stem (ES) cells. ES cell lines are derived from early-stage mouse embryos and can be maintained indefinitely in an undifferentiated state in vitro yet retain the capacity, when injected back into an early-stage mouse embryo, to mix with the endogenous cells of the embryo and contribute to all tissues of the developing mouse, including the germ line. The gene of interest is modified in ES cells by the introduction of a targeting vector that consists of a modified version of the endogenous gene. In ES cells, the targeting vector recombines with the homologous endogenous gene and thereby introduces the genetic modification. Gene targeted ES cells are then injected into wild type blastocyst-stage mouse embryos with the chimeric mice that result being mixtures of the modified ES cells and wild type blastocyst cells. The successful integration of the ES cells into the germ line permits the genetic modification to be propagated as part of the mouse genome, and this creates stable transgenic lines.

Gene targeting technology has been most commonly used to produce null mutations or gene knockouts. However, it can also be used to selectively modify endogenous mouse genes down to the level of introducing single nucleotide changes, producing what are known as knock-in mice. In contrast to pronuclear injection, in which multiple copies of a transgene are inserted randomly, in gene targeting, the normal mouse gene is modified in its normal chromosomal position. Thus, although in pronuclear injection a transgene is typically overexpressed and often misexpressed spatially and temporally because of its coupling to a heterologous promoter, a gene-targeted allele is expressed at normal levels and in a normal temporal and spatial expression pattern.

TRANSGENIC MODELING BASED ON OVEREXPRESSION OF THE AMYLOID PRECURSOR PROTEIN

Given the availability of methods for introducing genetic modifications, modeling in transgenic mice has been vigorously pursued on the basis of the amyloid hypothesis.[9–11] Initially, before the discovery of FAD mutations, attempts were made to overexpress wild-type APP in transgenic mice by pronuclear injection. Although a variety of promoters were tried, none of these efforts produced anything that resembled an amyloid plaque or any other recognizable AD-type pathology.

After the discovery of FAD mutations in APP, a number of groups turned their attention to making AD models based on the overexpression of transgenes containing FAD mutations. Games *et al.* 12 reported the first successful application of this approach using a platelet derived growth factor-β (PDGF) promoter to drive a human APP transgene that contained an FAD associated mutation (V717F). The PDGF promoter was chosen because, despite its name, it was known to be highly expressed in the central nervous system and to drive strong expression of exogenous transgenes in neurons. In the line that was generated (termed PDAPP because of the PDGF promoter plus APP), 40 copies of the transgene integrated, and this resulted in an approximately 18-fold elevation of APP RNA and an approximately 10-fold elevation of human APP protein in comparison with endogenous mouse APP levels. Proportionate increases in human Aβ were found.

PDAPP mice exhibited age-dependent amyloid deposition in the brain along with thioflavin-S–positive plaques, including compact plaques with dense cores that were highly reminiscent of those seen in human AD. Dystrophic neurites, reactive astrocytes, and activated microglia were all found near plaques. The process was age-related, in that plaque deposition was minimal at 6 months of age but readily apparent by 9 months, increasing dramatically by 12 to 15 months.[13] PDAPP mice were subsequently shown to develop age-related learning defects[14] and synapse loss.[15]

Independently, Hsiao *et al.*,[16] taking a relatively similar approach, overexpressed a human APP transgene containing the Swedish FAD mutation (K670N/M671L). Expression was driven by a hamster prion [prion protein (PrP)] promoter that drives expression widely in the nervous system. These mice, termed Tg2576 mice, expressed human APP at levels more than 5-fold above the levels of the endogenous mouse APP, and Aβ40 and Aβ42 levels increased with age. Like PDAPP mice, Tg2576 mice exhibited age-dependent amyloid deposition, which resulted in thioflavin-S–positive plaques similar to those found in AD, along with gliosis and dystrophic neurites. Plaque amyloid was first clearly seen by 11 to 13 months, eventually becoming widespread in cortical and limbic structures. Water maze learning, a test of spatial memory in mice, was normal in 3-month-old animals but impaired in 9- to 10-month-old mice. The Tg2576 mouse line has been made widely available and has been the most widely studied transgenic AD model.

Subsequently, many other transgenic lines were developed with approaches similar to those used to develop PDAPP and Tg2576 mice[9–11] typically relying on strong promoters to drive expression of APP transgenes containing single or multiple FAD mutations. Common features of the models have been the production of elevated levels of Aβ, amyloid plaques, dystrophic neurites, and gliosis. Behavioral deficits have been common.[10] Many additional neuropathological, electrophysiological, and neurochemical changes that model aspects of AD in humans have also been observed.[10]

Yet, although the models exhibit many similarities, they differ in certain aspects, one being the temporal profile of plaque deposition. For example, TgCRND8 mice, which express multiple APP mutations, exhibit parenchymal amyloid deposition as early as 3 months of age.[17] APP23 mice, in which a Thy-1 promoter is used to drive expression of the Swedish mutation, are notable for their prominent vascular amyloid deposition and resultant congophilic angiopathy.[18] PDAPP and Tg2576 mice also differ in some aspects; for example, in Tg2576 mice, both Aβ40 and Aβ42 are increased more or less proportionately, whereas in PDAPP mice, Aβ42 is disproportionately elevated.[16,19] In Tg2576 mice, most amyloid deposits in dense cored plaques, and Tg2576 mice have relatively few of the diffuse deposits found in PDAPP mice. Tg2576 mice are also known for their giant plaques[20] and exhibit more vascular amyloid deposition,[19] which is largely absent in PDAPP mice.

Variations between lines likely reflect the different promoters used, the effects of the distinct mutations or combinations of mutations, and the genetic backgrounds on which the transgenes have been maintained. PDAPP mice, for example, have been mostly studied on a highly mixed C57BL6/DBA/Swiss-Webster background. By contrast, Tg2576 mice are typically studied on a hybrid C57BL6/SJL background. Indeed, the Tg2576 transgene has been difficult to move off this background and leads to early death on an FVB/N background.[21]

TRANSGENIC MODELING WITH PRESENILIN MUTATIONS

Presenilin 1 (PS1) was found as part of a search for an early-onset FAD gene associated with a locus on chromosome 14.[8] Mutations in PS1 are the most commonly recognized causes of early-onset FAD, and to date, more than 160 mutations in PS1 linked to FAD have been discovered (see http://www.molgen.ua.ac.be/ADMutations). Mutations in a related gene on chromosome 1, now called *presenilin 2* (PS2), were soon linked to FAD as well.[8]

APP is a transmembrane protein with the C-terminus of the Aβ region located within the intramembranous domain.[22] Aβ is generated by the sequential cleavage of APP by a set of secretases. β-secretase cleaves APP at the N-terminus of the Aβ peptide, generating a C-terminal fragment that becomes the substrate for γ-secretase, which generates the Aβ C-terminus. By contrast, a third secretase, known as α-secretase, cleaves APP within the Aβ domain, precluding the generation of Aβ. Presenilins are best known for their role as part of the multiprotein complex known as γ-secretase, which executes the intramembranous cleavage of many transmembrane proteins, including APP.[23] Presenilin FAD mutations cause APP processing to be shunted toward the longer, more amyloidogenic Aβ42 species[23]; an observation that can be been seen as supportive of the amyloid hypothesis.

PS1 FAD mutant transgenic lines have been generated with many of the same promoters used to create APP mice, including PDGF[24] and PrP.[25,26] A few PS2 FAD mutant lines also exist. In addition, several gene-targeted lines exist in which PS1 FAD mutations have been targeted to the endogenous mouse PS1.[27–29] Presenilin FAD mutant mice consistently show elevations of Aβ42 with little if any effect on Aβ40. However, singly transgenic PS1 or PS2 mice do not develop plaques, although when crossed with plaque-forming APP lines, the presenilin FAD mutations cause earlier and more extensive plaque formation.[30] Why singly transgenic

Singly transgenic presenilin 1 or presenilin 2 mice do not develop plaques, although when crossed with plaque-forming amyloid precursor protein lines, the presenilin familial Alzheimer's disease mutations cause earlier and more extensive plaque formation.

PS1 and PS2 mice fail to develop plaque pathology is not entirely clear, but may be related to the generally lower levels of Aβ42 found in single presenilin transgenics versus APP-overexpressing lines as well as the lack of elevation of Aβ40 in presenilin transgenics. It may also be related to the differing aggregation properties of mouse Aβ versus human Aβ.[31]

Although PS1/APP bigenic mice have been frequently studied, the parental presenilin lines have been less studied, likely because of their lack of a robust AD-like pathology. However, PS1 and PS2 FAD mutant lines show exaggerated hippocampal damage after kainate-induced excitotoxicity,[27,32,33] and PS1 FAD mutants render animals more sensitive to trimethyltin-induced hippocampal damage.[34] Excessive neuronal loss in the entorhinal cortex also occurs in mice harboring the deltaE9 PS1 FAD mutation after lesioning of the perforant path.[35] Increased protein oxidation and lipid peroxidation have also been reported in PS1 FAD mutant brain.[36,37] Several studies have documented impaired hippocampal neurogenesis in adult PS1 FAD mutant mice,[38–42] and recently, an age-dependent impairment of spine morphology and synaptic plasticity in hippocampal cornu ammonis 1 neurons of a PS1 transgenic mouse model has been described.[43] Age-related neurodegenerative changes with neuronal loss have been reported in one PS1 FAD mutant line,[44] and age-related NFT-like inclusions have been described in a PS1 knock-in line.[45] Recently, a microvascular pathology that is highly reminiscent of the microvascular pathology found in AD has also been described.[46] Thus, presenilin FAD mutant mice exhibit a phenotype. What is less clear is why they fail to exhibit the full range of AD-related pathologies, given the potency of the mutations in humans.

RELEVANCE AS MODELS OF ALZHEIMER'S DISEASE

Transgenic mouse models now exist that mimic a range of AD-related pathologies. Table 1 contains a summary of some of the more widely studied mouse models. As discussed later, these models have suggested new insights into the pathophysiology as well as novel therapeutic approaches. The models,

Table 1. *Selected Examples of AD Mouse Models.*

Line	Promoter	FAD Mutation	Amyloid Pathology	Reference
PDAPP	PDGF	APP-Indiana	Parenchymal plaques at 6–9 months of age	12
Tg2576	PrP	APP-Swedish	Parenchymal plaques by 11–13 months of age with some vascular amyloid	16
APP23	Thy-1	APP-Swedish	Parenchymal plaques by 6 months of age and prominent vascular deposition of amyloid	18
TgCRND8	PrP	APP-Swedish + Indiana	More aggressive parenchymal plaque pathology present by 3 months of age	17
APP-Dutch	Thy-1	APPE693Q associated with hereditary cerebral hemorrhage with Dutch-type amyloidosis	Vascular deposition of amyloid with few parenchymal plaques	47
PS1M146V	PDGF	PS1M146V	Elevated Aβ42 without plaque pathology	24
PSAPP	PS1M146V × Tg2576	PS1M146V + APP-Swedish	Earlier and more extensive plaque pathology in comparison with Tg2576 alone	30
3xTg	Thy-1.2 and native mouse	Transgenes containing Thy-1.2–driven APP-Swedish and tau P301L were coinjected onto a homozygous PS1M146V knock-in background.	Parenchymal plaques by 6 months of age combined with tau pathology by 12 months of age	63

NOTE: Additional information on AD Tg models can be found at the Web site of the Alzheimer's Association (http://www.alzforum.org/res/com/tra).

Abbreviations: Aβ, β-amyloid; AD, Alzheimer's disease; APP, amyloid precursor protein; FAD, familial Alzheimer's disease; PDAPP, platelet-derived growth factor promoter driving amyloid precursor protein; PDGF, platelet-derived growth factor β; PrP, prion protein; PS1, presenilin 1; PSAPP, presenilin/amyloid precursor protein; Tg, transgenic.

however, raise a number of issues as well. First, it is clear that the success of transgenic mouse models has depended on the overexpression of APP transgenes containing FAD-associated mutations at levels that are not physiological. Indeed, many proteins, if overexpressed at sufficient levels, will become toxic at some point. It might be argued, therefore, that it is hardly surprising that overexpressing a naturally amyloidogenic protein would lead to amyloid deposits. However, interestingly, even though a number of APP wild transgenic lines have been created, only 1 transgenic mouse line overexpressing wild-type APP has been described that develops plaques.[47] Thus, in general, overexpression of wild-type APP in the mouse does not induce plaque pathology, and the pathology that is seen seems to require the presence of an FAD mutation.

It is clear that the success of transgenic mouse models has depended on the overexpression of APP transgenes containing FAD-associated mutations at levels that are not physiological.

It is interesting in this context to compare the overexpression models to APP knock-in mice harboring the Swedish mutation introduced into the mouse APP gene.[48] These mice should represent the most authentic model of human FAD in the mouse, and because the 3 amino acid differences between mouse and human APPs in the Aβ region were modified, these mice also produce human Aβ. Yet, despite increased Aβ production, no amyloid plaques or neuropathology develops in these mice.

Overexpression of APP in Down's syndrome is associated with AD pathology in humans.[49] Recently, families have been identified in which a duplication of the APP gene leads to FAD, presumably because of overexpression of wild-type APP.[50,51] Thus, there is a precedent for APP overexpression leading to AD in humans. Yet, there is no evidence for APP overexpression in either sporadic AD or in FAD associated with APP or presenilin mutations.

It is also not clear that the temporal appearance of plaque pathology and cognitive deficits in the mouse models mirrors that of the human disease. Indeed, many of the mouse models exhibit behavioral deficits before significant plaque pathology.[10] By contrast, substantial plaque pathology is probably present for some period of time before cognitive symptoms first appear in human AD.[52] Interpretations of behavioral studies in the mouse are also complicated by the fact that it is difficult to know how precisely the behavioral tests used in mice model the cognitive deficits in humans.

Transgenic mouse models have also been troubled by the difficulty of producing the full spectrum of AD pathologies, including neuronal loss. For example, neither PDAPP nor Tg2576 mice, despite having extensive amyloid deposition, exhibit neuronal loss.[53,54] APP23 mice show only modest losses of cornu ammonis 1 pyramidal cells (about 15%),[55] losses that are far less than those observed in AD. More substantial neuronal loss has been reported in mice expressing multiple PS1 and APP mutations.[56–58] For example, mice harboring 2 APP and 1 PS1 FAD mutation show a 35% loss of hippocampal neurons.[57] In another study, 2 APP mutations (APP-Swedish/APP-London) expressed from transgenes were crossed onto a mouse line that had 2 presenilin FAD mutations (M233T/L235P) knocked into the mouse PS1 gene. Besides amyloid pathology, extensive neuronal loss (>50%) occurred in the hippocampus.[56] Finally, 3 APP and 2 PS1 FAD mutations recently were combined to create a 5X FAD mutant mouse, and grossly apparent neuronal loss was observed.[58] Thus, neuronal loss can be induced in the mouse but only, it would seem, by the combination of multiple mutations that individually are sufficient to cause disease in humans.

Transgenic mouse models have also been troubled by the difficulty of producing the full spectrum of AD pathologies, including neuronal loss.

Another problem with the models has been the general difficulty of inducing the characteristic cytoskeletal pathology of AD in transgenic mice.[59] NFTs are recognized by their propensity to stain with histological dyes, such as thioflavin-S and Congo red, and by ultrastructurally containing paired helical filaments. NFTs contain hyperphosphorylated forms of tau as well as conformationally altered tau epitopes that can be recognized by specific antibodies. In PDAPP mice, phosphorylated tau epitopes do accumulate within dystrophic neurites in animals 14 months of age or older,[60] and within these neurites, 12- to 15-nm filaments can be seen, but there are no paired helical filaments and no lesions with the histological staining properties of NFTs. Other models have been similar in their lack of any NFT-like pathology. Various suggestions have been advanced for the difficulty of inducing NFT-like lesions in the mouse, including differences between human and mouse tau and the shorter life span of the mouse. However, the reasons remain unexplained.

Recently, mice that exhibit NFT-like lesions and plaques have been produced by combining FAD mutations with mutant forms of tau found in a distinct form of dementia known as frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17).[61] FTDP-17 is part of the spectrum of frontotemporal dementias (FTDs) that include Pick's disease, progressive supranuclear palsy, and corticobasal degeneration.[61] In these disorders, hyperphosphorylated forms of tau accumulate in NFTs in the absence of neuritic plaques. FTD cases present clinically with often prominent behavioral disturbances or isolated language defects and relatively preserved memory initially. Like AD, most FTD occurs sporadically. However, in some families with FTDP, the disease is caused by mutations in the tau gene on chromosome 17 (FTDP-17). Mutations, including exonic point mutations, which alter the protein coding sequence, and intronic mutations, which affect splicing or the level of tau expression, have all been identified.

Recently, mice that exhibit neurofibrilary-tangle-like lesions and plaques have been produced by combining FAD mutations with mutant forms of tau found in a distinct form of dementia known as frontotemporal dementia and parkinsonism linked to chromosome 17 (FTDP-17).

Lewis *et al.*[62] first crossed a transgenic line known as JPNL3, which expresses the P301L mutation associated with FTDP-17, with Tg2576 mice. Singly transgenic JPNL3 mice were known to develop NFT-like lesions, and bigenic mice [tau amyloid precursor protein (TAPP) mice] exhibited NFTs and amyloid plaques. More recently, Oddo *et al.*[63] generated a triply transgenic model (3xTg-AD) by coinjecting 2 transgenes containing APP-Swedish and P301L FTDP-17 mutations into fertilized eggs harvested from PS1M146V knock-in mice. The resulting transgenic lines thus expressed mutations in APP and tau from exogenous transgenes combined with a PS1 FAD mutation from the endogenous mouse gene. With aging, these mice had increased Aβ40 and Aβ42 levels. Intraneuronal Aβ accumulated, and the mice exhibited amyloid plaques and NFT-like lesions that immunostained with antibodies recognizing conformationally altered tau epitopes. Amyloid plaques were apparent by 6 months of age and preceded tau pathology, which was not evident until about 1 year of age. Functionally, 3xTg-AD mice developed age-dependent synaptic dysfunction, including altered long-term potentiation and deficits in spatial memory.[63,64]

TAPP and 3xTg-AD thus exhibit a broader spectrum of AD pathologies. The question, however, again arises: are these models of AD? They do exhibit the much sought after plaques with a tangle pathology. However, they represent a composite of 2 distinct diseases that do not naturally occur together. Plaque development is almost certainly driven by the APP and PS1 FAD mutations, whereas the tangle-like pathology is driven by the tau mutations. Indeed, in 3xTg-AD mice, the pathologies arise in spatially distinct patterns, with Aβ deposition starting in the neocortex and appearing later in the hippocampus, whereas tau pathology progresses in the opposite direction, beginning in the hippocampus and appearing later in the neocortex.[63] It does, however, appear that the mutations interact because in TAPP mice, NFTs are found in regions that rarely or never exhibit NFTs in singly transgenic JPNL3 animals. In addition, intracerebral injections of anti-Aβ antibodies into the hippocampus of 3xTg-AD mice not only reduced Aβ accumulation but also resulted in clearance of early-stage tau pathology, although later-stage lesions were resistant.[65] These studies thus show that modulating Aβ affects tau pathology and suggest that tau pathology may be downstream of Aβ generation; this notion has also been supported by recent studies showing that behavioral deficits in APP FAD mutant mice are reduced when endogenous tau is removed by breeding the APP transgene onto a tau-null background.[66]

Finally, all the transgenic mouse models discussed are based on the expression of mutant genes implicated in causing AD in a relatively small subset of patients. Except for their generally early age of onset, these familial cases appear to phenocopy quite well the clinical and pathological features of sporadic cases, but it remains possible that the mutations themselves introduce effects that are not found in the sporadic disease. In particular, presenilins affect a range of biological functions through effects that are not dependent on their role in γ-secretase activity, and at least some of these activities are known to be altered by FAD mutants.[23]

USE OF TRANSGENIC MODELS TO STUDY THE PATHOPHYSIOLOGY OF ALZHEIMER'S DISEASE: EFFECTS OF DIFFERENT β-AMYLOID SPECIES

Transgenic mice have been used to study many aspects of AD pathophysiology. Two recent examples of their usefulness are studies designed to test differences in the in vivo effects of Aβ40 and Aβ42. In vitro, Aβ42 is more amyloidogenic than Aβ40.[67] However, Aβ40 is found in plaque amyloid and is especially common in vascular amyloid.[68] In vitro studies have suggested, however, that Aβ42 deposition is necessary in order to seed Aβ40-containing deposits.[69] To test the relative plaque-forming propensity in vivo, transgenic mice were generated that selectively express either Aβ40 or Aβ42.[70] These mice were created by the fusion of Aβ40 and Aβ42 coding sequences to the C-terminus of the BRI protein associated with familial British dementia. BRI is a transmembrane protein that undergoes constitutive cleavage near its C-terminus that releases a soluble 23–amino acid peptide. Aβ40 or Aβ42 selective expression systems were created by the replacement of the 23–amino acid BRI peptide region with human Aβ40 or Aβ42 sequences. Thus, in the fusion construct, cleavage, rather than releasing the native BRI peptide, releases either human Aβ40 or Aβ42. Although highly artificial constructs, the Bri-Aβ40 and Bri-Aβ42 transgenes offer an elegantly simple system for introducing human Aβ40 or Aβ42 into the mouse brain. Insoluble Aβ42 accumulated with aging in BRI-Aβ42 transgenic mice and was accompanied by amyloid deposition in the parenchyma in the form of both plaques and diffuse deposits as well as extensive vascular amyloid deposition. By contrast, BRI-Aβ40 mice developed no amyloid pathology at any age. These studies are thus consistent with Aβ42 being required for parenchymal and vascular deposits of Aβ40, and indeed, more recent studies have shown that BRI mice expressing high levels of Aβ40 are protected against amyloid deposition when bred with BRI-Aβ42 mice.[71]

Both parenchymal and vascular deposits of amyloid are typical of AD. There are, however, mutations in APP that do not produce substantial parenchymal deposition in the form of plaques but result in extensive vascular deposits. These mutations, rather than producing FAD, cause a hereditary form of recurrent intracerebral hemorrhage. One such mutation, E693Q, is known as the Dutch mutation. Interestingly, although most FAD mutations in APP reside near the β- and γ-secretase cleavage sites, mutations associated with cerebral hemorrhage such as the Dutch mutation tend to be found within the Aβ domain itself.

Herzig *et al.*[47] developed transgenic lines that expressed human wild-type APP, APP-Dutch, or APP-Dutch crossed with a PS1 FAD mutant. These lines exhibited various Aβ40/Aβ42 ratios, with the APP-Dutch ratio higher than the wild-type APP ratio, which was in turn greater than the APP-Dutch/PS1 ratio. The high Aβ40/Aβ42 ratio in the APP-Dutch mice resulted in extensive vascular amyloid deposition with essentially no parenchymal deposition. By contrast, APP-Dutch/PS1 mice with about half the Aβ40/Aβ42 ratio of APP-Dutch mice developed parenchymal plaques with little vascular deposition, and wild-type APP mice with intermediate Aβ40/Aβ42 ratios had mixed parenchymal and vascular deposition. Thus, although Aβ42 may be needed as a seed for amyloid deposition in either compartment, these studies suggest that Aβ40 promotes vascular deposition, whereas Aβ42 shifts deposition toward parenchymal amyloid. Other studies[72] have, however, found that transgenic mice harboring a distinct 693 mutation (E693G/Arctic) combined with APP-Swedish and APP-Indiana mutations develop prominent parenchymal plaque deposits with little congophilic angiopathy despite high Aβ40/Aβ42 ratios. Thus, some property of mutations at the 693 site, besides their effect on Aβ40/Aβ42 ratios, also appears to influence parenchymal deposition versus vascular deposition.

TRANSGENIC MODELS AND THE RISE OF THE β-AMYLOID OLIGOMER

The original amyloid hypothesis regarded Aβ deposited in plaque amyloid as the toxic species. However, subsequently it became clear that plaque counts correlate relatively poorly with the level of cognitive decline and that in fact the number of NFTs correlates more strongly with the degree of dementia.[73] This led to a reevaluation of what form of amyloid might constitute the most toxic species, and recently, soluble forms of Aβ have been proposed as the more toxic species.

Soluble Aβ species are toxic in cell culture.[74] APP transgenic mice have provided strong circumstantial evidence for the toxicity of Aβ oligomers in vivo by showing that many pathological and functional changes in mice occur before the appearance of plaque pathology. For example, a series of studies in PDAPP mice demonstrated that volume loss and other anatomic changes in the dentate gyrus are

DOI:10.1002/MSJ

APP transgenic mice have provided strong circumstantial evidence for the toxicity of Aβ oligomers in vivo by showing that many pathological and functional changes in mice occur before the appearance of plaque pathology.

present in 100-day-old mice well before plaque deposition.[75,76] Tg2576 mice also exhibit electrophysiological and behavioral changes months before plaque deposition,[77] and axonal swellings may be present in Tg2576 mice for up to a year before amyloid deposition.[78] Behavioral and electrophysiological changes have been described in other lines of APP mice before amyloid deposition as well.[30,79,80] The finding of pathophysiological changes in PS1 FAD mutant mice (discussed previously) that have elevated Aβ42 but no plaques is also consistent with a toxic role for soluble forms of Aβ.

Recently, a 56-kDa oligomeric Aβ species identified in Tg2576 mice[81] has been proposed as a candidate for the toxic effect. Levels of this species, which has been termed Aβ*56, correlate strongly with the degree of memory impairment in Tg2576 mice, and when injected into rats, Aβ*56 disrupts cognitive functioning.[81] Studies in 3xTg-AD suggest that levels of Aβ*56 may correlate with pathology and behavioral changes in this line as well.[82]

β-AMYLOID IMMUNIZATION AS A THERAPEUTIC STRATEGY IN ALZHEIMER'S DISEASE

Transgenic animals have been widely used to study factors that affect Aβ deposition. For example, the role that apolipoprotein ε4 and its isoforms play in the development of plaque pathology has been extensively studied with apolipoprotein ε4 transgenic and knockout mice crossed to APP transgenic mice.[83–85] Studies examining how dietary cholesterol affects plaque pathology in transgenic mice[86] have formed part of the basis for proposals to use statin therapy for AD. Indeed, APP transgenic mice have been used to evaluate a variety of factors that influence Aβ production and plaque pathology; these factors range from inflammatory modulators, metal chelators, and natural products that bind Aβ to lifestyle factors including exercise, environmental enrichment, caloric restriction, and wine consumption.[10,87–90]

However, nowhere have transgenic mice played a larger role than in the development of potential immunotherapies for AD.

However, nowhere have transgenic mice played a larger role than in the development of potential immunotherapies for AD. The initial impetus for immunotherapeutic approaches in AD came from in vitro observations showing that anti-Aβ antibodies could prevent fibril formation and disaggregate preformed fibrils.[91] However, without studies in transgenic mice demonstrating potential effectiveness in vivo, it seems unlikely that these approaches would have ever made their way into human trials.

Schenk *et al.*,[92] using aggregated Aβ as an immunogen, first investigated Aβ vaccination in PDAPP mice. They started immunizing mice at 6 weeks of age (well before the age when plaque deposition begins in this line) with a vaccine that was injected on a monthly basis for 11 months. High titers of anti-Aβ antibodies developed, and at 12 months of age, the mice showed dramatic reductions in plaque loads in comparison with saline-injected controls. In an additional set of studies, 11-month-old mice with already existing plaque loads received monthly injections and were examined pathologically at 15 and 18 months. At both ages, plaque burdens were dramatically less in the immunized animals, suggesting that Aβ vaccination might be beneficial even after parenchymal deposits are present.

Other groups soon confirmed the beneficial effects of active immunization in other transgenic lines and showed that vaccination can reverse spatial memory deficits in APP and presenilin/APP mice.[93,94] Strikingly, behavioral performance often improved more than brain Aβ levels or plaque pathology, with deficits in some lines being almost entirely normalized.[17] Subsequently, it was found that passive immunization of anti-Aβ antibodies peripherally also reduced amyloid deposition.[95,96] Indeed, multiple studies have shown that even short-term administration of anti-Aβ antibodies improves performance in tests of learning and memory.[97] Improvements in behavioral performance with passive immunization strategies also occur even when no detectible effect is seen on deposits of plaque amyloid. For example, Dodart *et al.*[98] gave weekly injections of the anti-Aβ monoclonal antibody m266 for 6 weeks to 24-month-old PDAPP mice and found that behavioral impairments were essentially normalized despite there being no effects on plaque pathology. Indeed, even a single injection of this antibody to 11-month-old mice produced beneficial effects on behavior that were apparent after 1 day of treatment.[98] Similar effects have been observed with a different monoclonal antibody in Tg2576 mice.[99]

How immunization works remains uncertain. Multiple mechanisms have been suggested, including stimulation of antibody-induced phagocytosis of plaque deposits as well as induction of cell-mediated immune responses. Anti-Aβ antibodies might also directly disrupt existing aggregates, impede the formation of new aggregates, or neutralize toxic soluble species of Aβ. Peripheral administration of anti-Aβ antibodies has also been shown to induce massive increases in plasma Aβ,[96,100] suggesting that passive immunization may promote an efflux of Aβ out of the brain into what has been called a peripheral sink. These mechanisms as well as others have recently been reviewed.[97,101]

Whatever its mechanism of action, studies in transgenic mice were sufficiently compelling that Aβ immunization was taken into clinical trials in humans. In the first of these studies, the AN1792 trial, subjects were injected with an aggregated Aβ preparation similar to that used in the initial mouse studies. Although phase 1 studies indicated adequate immunological responses and tolerability of the vaccine, phase 2 studies in patients with AD were halted because of the appearance of an autoimmune meningoencephalitis in some patients.[102] Although these initial studies were ultimately disappointing, a retrospective analysis found that those patients with the highest anti-Aβ titers had significantly less cognitive decline than those with low titers,[103] arguing that active immunization remains a viable strategy if autoimmune side effects can be overcome. Most human trials are, however, currently pursuing passive immunization strategies because of the greater control over antibody levels that can be achieved as well as concerns over the ability of older subjects to mount an effective immune response. The farthest advanced of these trails at the time of writing is the Elan/Wyeth trial of AAB-001 (bapineuzumab), which entered phase III testing in 2007 and uses a humanized monoclonal antibody directed against the N-terminus of Aβ. More information on this and other current clinical trials can be found at http://www.clinicaltrials.gov and http://alzforum.com.

CONCLUSION

Animal models attempt to mimic human diseases and thus allow testing of therapeutic strategies as well as the investigation of the disease course and underlying pathophysiology in a manner that is impractical or unethical in humans. Transgenic models now dominate approaches to animal modeling of human neurodegenerative diseases, including AD. In one sense, the models are limited by their requirement for a genetic modification. However, the discovery of genes for familial forms of AD has allowed transgenic models to be created that reproduce many critical aspects of the disease. These models have improved our understanding of the disease pathogenesis and have led to new therapeutic approaches. They will, without a doubt, continue to play central roles for years to come in preclinical testing and be used as tools for developing insight into the biological basis of AD.

ACKNOWLEDGMENT

Work conducted in the authors' laboratory has been supported by grants from the National Institute on Aging (AG20139, AG029361, AG02219, and AG05138), the Alzheimer's Association (IIRG-07-57318) and a Merit Award from the Department of Veterans Affairs (1I01BX000342).

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Link CD. Invertebrate models of Alzheimer's disease. *Genes Brain Behav* 2005; 4: 147–156.
2. Wu Y, Luo Y. Transgenic *C. elegans* as a model in Alzheimer's research. *Curr Alzheimer Res* 2005; 2: 37–45.
3. Folkesson R, Malkiewicz K, Kloskowska E, et al. A transgenic rat expressing human APP with the Swedish Alzheimer's disease mutation. *Biochem Biophys Res Commun* 2007; 358: 777–782.
4. Agca C, Fritz JJ, Walker LC, et al. Development of transgenic rats producing human beta-amyloid precursor protein as a model for Alzheimer's disease: transgene and endogenous APP genes are regulated tissue-specifically. *BMC Neurosci* 2008; 9: 28.
5. Flood DG, Lin YG, Lang DM, et al. A transgenic rat model of Alzheimer's disease with extracellular Abeta deposition. *Neurobiol Aging* 2009; 30: 1078–1090.
6. Liu L, Orozco IJ, Planel E, et al. A transgenic rat that develops Alzheimer's disease-like amyloid

pathology, deficits in synaptic plasticity and cognitive impairment. *Neurobiol Dis* 2008; 31: 46–57.

7. Selkoe DJ. Alzheimer's disease: genes, proteins, and therapy. *Physiol Rev* 2001; 81: 741–766.
8. Ertekin-Taner N. Genetics of Alzheimer's disease: a centennial review. *Neurol Clin* 2007; 25: 611–667.
9. Codita A, Winblad B, Mohammed AH. Of mice and men: more neurobiology in dementia. *Curr Opin Psychiatry* 2006; 19: 555–563.
10. Games D, Buttini M, Kobayashi D, et al. Mice as models: transgenic approaches and Alzheimer's disease. *J Alzheimers Dis* 2006; 9(suppl): 133–149.
11. McGowan E, Eriksen J, Hutton M. A decade of modeling Alzheimer's disease in transgenic mice. *Trends Genet* 2006; 22: 281–289.
12. Games D, Adams D, Alessandrini R, et al. Alzheimer-type neuropathology in transgenic mice overexpressing V717F beta-amyloid precursor protein. *Nature* 1995; 373: 523–527.
13. Reilly JF, Games D, Rydel RE, et al. Amyloid deposition in the hippocampus and entorhinal cortex: quantitative analysis of a transgenic mouse model. *Proc Natl Acad Sci U S A* 2003; 100: 4837–4842.
14. Chen G, Chen KS, Knox J, et al. A learning deficit related to age and beta-amyloid plaques in a mouse model of Alzheimer's disease. *Nature* 2000; 408: 975–979.
15. Dodart JC, Mathis C, Saura J, et al. Neuroanatomical abnormalities in behaviorally characterized APP(V717F) transgenic mice. *Neurobiol Dis* 2000; 7: 71–85.
16. Hsiao K, Chapman P, Nilsen S, et al. Correlative memory deficits, Abeta elevation, and amyloid plaques in transgenic mice. *Science* 1996; 274: 99–102.
17. Chishti MA, Yang DS, Janus C, et al. Early-onset amyloid deposition and cognitive deficits in transgenic mice expressing a double mutant form of amyloid precursor protein 695. *J Biol Chem* 2001; 276: 21562–21570.
18. Calhoun ME, Burgermeister P, Phinney AL, et al. Neuronal overexpression of mutant amyloid precursor protein results in prominent deposition of cerebrovascular amyloid. *Proc Natl Acad Sci U S A* 1999; 96: 14088–14093.
19. Fryer JD, Taylor JW, DeMattos RB, et al. Apolipoprotein E markedly facilitates age-dependent cerebral amyloid angiopathy and spontaneous hemorrhage in amyloid precursor protein transgenic mice. *J Neurosci* 2003; 23: 7889–7896.
20. Sasaki A, Shoji M, Harigaya Y, et al. Amyloid cored plaques in Tg2576 transgenic mice are characterized by giant plaques, slightly activated microglia, and the lack of paired helical filament-typed, dystrophic neurites. *Virchows Arch* 2002; 441: 358–367.
21. Hsiao KK, Borchelt DR, Olson K, et al. Age-related CNS disorder and early death in transgenic FVB/N mice overexpressing Alzheimer amyloid precursor proteins. *Neuron* 1995; 15: 1203–1218.
22. Zheng H, Koo EH. The amyloid precursor protein: beyond amyloid. *Mol Neurodegener* 2006; 1: 5.
23. Vetrivel KS, Zhang YW, Xu H, Thinakaran G. Pathological and physiological functions of presenilins. *Mol Neurodegener* 2006; 1: 4.
24. Duff K, Eckman C, Zehr C, et al. Increased amyloid-beta42(43) in brains of mice expressing mutant presenilin 1. *Nature* 1996; 383: 710–713.
25. Borchelt DR, Thinakaran G, Eckman CB, et al. Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1-42/1-40 ratio in vitro and in vivo. *Neuron* 1996; 17: 1005–1013.
26. Citron M, Westaway D, Xia W, et al. Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. *Nat Med* 1997; 3: 67–72.
27. Guo Q, Sebastian L, Sopher BL, et al. Increased vulnerability of hippocampal neurons from presenilin-1 mutant knock-in mice to amyloid beta-peptide toxicity: central roles of superoxide production and caspase activation. *J Neurochem* 1999; 72: 1019–1029.
28. Flood DG, Reaume AG, Dorfman KS, et al. FAD mutant PS-1 gene-targeted mice: increased A beta 42 and A beta deposition without APP overproduction. *Neurobiol Aging* 2002; 23: 335–348.
29. Nakano Y, Kondoh G, Kudo T, et al. Accumulation of murine amyloidbeta42 in a gene-dosage-dependent manner in PS1 'knock-in' mice. *Eur J Neurosci* 1999; 11: 2577–2581.
30. Holcomb L, Gordon MN, McGowan E, et al. Accelerated Alzheimer-type phenotype in transgenic mice carrying both mutant amyloid precursor protein and presenilin 1 transgenes. *Nat Med* 1998; 4: 97–100.
31. Jankowsky JL, Younkin LH, Gonzales V, et al. Rodent A beta modulates the solubility and distribution of amyloid deposits in transgenic mice. *J Biol Chem* 2007; 282: 22707–22720.
32. Grilli M, Diodato E, Lozza G, et al. Presenilin-1 regulates the neuronal threshold to excitotoxicity both physiologically and pathologically. *Proc Natl Acad Sci U S A* 2000; 97: 12822–12827.
33. Schneider I, Reverse D, Dewachter I, et al. Mutant presenilins disturb neuronal calcium homeostasis in the brain of transgenic mice, decreasing the threshold for excitotoxicity and facilitating long-term potentiation. *J Biol Chem* 2001; 276: 11539–11544.
34. Kassed CA, Butler TL, Navidomskis MT, et al. Mice expressing human mutant presenilin-1 exhibit decreased activation of NF-kappaB p50 in hippocampal neurons after injury. *Mol Brain Res* 2003; 110: 152–157.
35. Lazarov O, Peterson LD, Peterson DA, Sisodia SS. Expression of a familial Alzheimer's disease-linked presenilin-1 variant enhances perforant pathway lesion-induced neuronal loss in the entorhinal cortex. *J Neurosci* 2006; 26: 429–434.
36. Mohmmad Abdul H, Wenk GL, Gramling M, et al. APP and PS-1 mutations induce brain oxidative stress independent of dietary cholesterol: implications for Alzheimer's disease. *Neurosci Lett* 2004; 368: 148–150.
37. Schuessel K, Frey C, Jourdan C, et al. Aging sensitizes toward ROS formation and lipid peroxidation in PS1M146L transgenic mice. *Free Radic Biol Med* 2006; 40: 850–862.
38. Chevallier NL, Soriano S, Kang DE, et al. Perturbed neurogenesis in the adult hippocampus associated with presenilin-1 A246E mutation. *Am J Pathol* 2005; 167: 151–159.
39. Choi SH, Veeraraghavalu K, Lazarov O, et al. Non-cell-autonomous effects of presenilin 1 variants on

enrichment-mediated hippocampal progenitor cell proliferation and differentiation. *Neuron* 2008; 59: 568–580.

40. Wang R, Dineley KT, Sweatt JD, Zheng H. Presenilin 1 familial Alzheimer's disease mutation leads to defective associative learning and impaired adult neurogenesis. *Neuroscience* 2004; 126: 305–312.
41. Wen PH, Shao X, Shao Z, et al. Overexpression of wild type but not an FAD mutant presenilin-1 promotes neurogenesis in the hippocampus of adult mice. *Neurobiol Dis* 2002; 10: 8–19.
42. Wen PH, Hof PR, Chen X, et al. The presenilin-1 familial Alzheimer disease mutant P117L impairs neurogenesis in the hippocampus of adult mice. *Exp Neurol* 2004; 188: 224–237.
43. Auffret A, Gautheron V, Repici M, et al. Age-dependent impairment of spine morphology and synaptic plasticity in hippocampal CA1 neurons of a presenilin 1 transgenic mouse model of Alzheimer's disease. *J Neurosci* 2009; 29: 10144–10152.
44. Chui DH, Tanahashi H, Ozawa K, et al. Transgenic mice with Alzheimer presenilin 1 mutations show accelerated neurodegeneration without amyloid plaque formation. *Nat Med* 1999; 5: 560–564.
45. Tanemura K, Chui DH, Fukuda T, et al. Formation of tau inclusions in knock-in mice with familial Alzheimer disease (FAD) mutation of presenilin 1 (PS1). *J Biol Chem* 2006; 281: 5037–5041.
46. Gama Sosa M, De Gasperi R, Rocher A, et al. Age-related vascular pathology in transgenic mice expressing presenilin 1-associated familial Alzheimer's disease mutations. *Am J Pathol.* DOI:10.2353/ajpath.2010.090482.
47. Herzig MC, Winkler DT, Burgermeister P, et al. Abeta is targeted to the vasculature in a mouse model of hereditary cerebral hemorrhage with amyloidosis. *Nat Neurosci* 2004; 7: 954–960.
48. Reaume AG, Howland DS, Trusko SP, et al. Enhanced amyloidogenic processing of the beta-amyloid precursor protein in gene-targeted mice bearing the Swedish familial Alzheimer's disease mutations and a "humanized" Abeta sequence. *J Biol Chem* 1996; 271: 23380–23388.
49. Zigman WB, Schupf N, Sersen E, Silverman W. Prevalence of dementia in adults with and without Down syndrome. *Am J Ment Retard* 1996; 100: 403–412.
50. Kasuga K, Shimohata T, Nishimura A, et al. Identification of independent APP locus duplication in Japanese patients with early-onset Alzheimer disease. *J Neurol Neurosurg Psychiatry* 2009; 80: 1050–1052.
51. Sleegers K, Brouwers N, Gijselinck I, et al. APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. *Brain* 2006; 129: 2977–2983.
52. Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. *Ann Neurol* 1999; 45: 358–368.
53. Irizarry MC, McNamara M, Fedorchak K, et al. APPSw transgenic mice develop age-related A beta deposits and neuropil abnormalities, but no neuronal loss in CA1. *J Neuropathol Exp Neurol* 1997; 56: 965–973.
54. Irizarry MC, Soriano F, McNamara M, et al. Abeta deposition is associated with neuropil changes, but not with overt neuronal loss in the human amyloid precursor protein V717F (PDAPP) transgenic mouse. *J Neurosci* 1997; 17: 7053–7059.
55. Calhoun ME, Wiederhold KH, Abramowski D, et al. Neuron loss in APP transgenic mice. *Nature* 1998; 395: 755–756.
56. Casas C, Sergeant N, Itier JM, et al. Massive CA1/2 neuronal loss with intraneuronal and N-terminal truncated Abeta42 accumulation in a novel Alzheimer transgenic model. *Am J Pathol* 2004; 165: 1289–1300.
57. Schmitz C, Rutten BP, Pielen A, et al. Hippocampal neuron loss exceeds amyloid plaque load in a transgenic mouse model of Alzheimer's disease. *Am J Pathol* 2004; 164: 1495–1502.
58. Oakley H, Cole SL, Logan S, et al. Intraneuronal beta-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: potential factors in amyloid plaque formation. *J Neurosci* 2006; 26: 10129–10140.
59. Duyckaerts C, Potier MC, Delatour B. Alzheimer disease models and human neuropathology: similarities and differences. *Acta Neuropathol* 2008; 115: 5–38.
60. Masliah E, Sisk A, Mallory M, Games D. Neurofibrillary pathology in transgenic mice overexpressing V717F beta-amyloid precursor protein. *J Neuropathol Exp Neurol* 2001; 60: 357–368.
61. Haugarvoll K, Wszolek ZK, Hutton M. The genetics of frontotemporal dementia. *Neurol Clin* 2007; 25: 697–715.
62. Lewis J, Dickson DW, Lin WL, et al. Enhanced neurofibrillary degeneration in transgenic mice expressing mutant tau and APP. *Science* 2001; 293: 1487–1491.
63. Oddo S, Caccamo A, Shepherd JD, et al. Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular Abeta and synaptic dysfunction. *Neuron* 2003; 39: 409–421.
64. Billings LM, Oddo S, Green KN, et al. Intraneuronal Abeta causes the onset of early Alzheimer's disease-related cognitive deficits in transgenic mice. *Neuron* 2005; 45: 675–688.
65. Oddo S, Billings L, Kesslak JP, et al. Abeta immunotherapy leads to clearance of early, but not late, hyperphosphorylated tau aggregates via the proteasome. *Neuron* 2004; 43: 321–332.
66. Roberson ED, Scearce-Levie K, Palop JJ, et al. Reducing endogenous tau ameliorates amyloid beta-induced deficits in an Alzheimer's disease mouse model. *Science* 2007; 316: 750–754.
67. Barrow CJ, Yasuda A, Kenny PT, Zagorski MG. Solution conformations and aggregational properties of synthetic amyloid beta-peptides of Alzheimer's disease. Analysis of circular dichroism spectra. *J Mol Biol* 1992; 225: 1075–1093.
68. Kumar-Singh S. Cerebral amyloid angiopathy: pathogenetic mechanisms and link to dense amyloid plaques. *Genes Brain Behav* 2008; 7(suppl 1): 67–82.
69. Fryer JD, Holtzman DM. The bad seed in Alzheimer's disease. *Neuron* 2005; 47: 167–168.
70. McGowan E, Pickford F, Kim J, et al. Abeta42 is essential for parenchymal and vascular amyloid deposition in mice. *Neuron* 2005; 47: 191–199.
71. Kim J, Onstead L, Randle S, et al. Abeta40 inhibits amyloid deposition in vivo. *J Neurosci* 2007; 27: 627–633.

72. Cheng IH, Palop JJ, Esposito LA, et al. Aggressive amyloidosis in mice expressing human amyloid peptides with the Arctic mutation. *Nat Med* 2004; 10: 1190–1192.
73. Arriagada PV, Growdon JH, Hedley-Whyte ET, Hyman BT. Neurofibrillary tangles but not senile plaques parallel duration and severity of Alzheimer's disease. *Neurology* 1992; 42(pt 1): 631–639.
74. Walsh DM, Selkoe DJ. A beta oligomers–a decade of discovery. *J Neurochem* 2007; 101: 1172–1184.
75. Redwine JM, Kosofsky B, Jacobs RE, et al. Dentate gyrus volume is reduced before onset of plaque formation in PDAPP mice: a magnetic resonance microscopy and stereologic analysis. *Proc Natl Acad Sci U S A* 2003; 100: 1381–1386.
76. Wu CC, Chawla F, Games D, et al. Selective vulnerability of dentate granule cells prior to amyloid deposition in PDAPP mice: digital morphometric analyses. *Proc Natl Acad Sci U S A* 2004; 101: 7141–7146.
77. Jacobsen JS, Wu CC, Redwine JM, et al. Early-onset behavioral and synaptic deficits in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci U S A* 2006; 103: 5161–5166.
78. Stokin GB, Lillo C, Falzone TL, et al. Axonopathy and transport deficits early in the pathogenesis of Alzheimer's disease. *Science* 2005; 307: 1282–1288.
79. Hsia AY, Masliah E, McConlogue L, et al. Plaque-independent disruption of neural circuits in Alzheimer's disease mouse models. *Proc Natl Acad Sci U S A* 1999; 96: 3228–3233.
80. Moechars D, Dewachter I, Lorent K, et al. Early phenotypic changes in transgenic mice that overexpress different mutants of amyloid precursor protein in brain. *J Biol Chem* 1999; 274: 6483–6492.
81. Lesne S, Koh MT, Kotilinek L, et al. A specific amyloid-beta protein assembly in the brain impairs memory. *Nature* 2006; 440: 352–357.
82. Billings LM, Green KN, McGaugh JL, LaFerla FM. Learning decreases A beta*56 and tau pathology and ameliorates behavioral decline in 3xTg-AD mice. *J Neurosci* 2007; 27: 751–761.
83. Bales KR, Verina T, Dodel RC, et al. Lack of apolipoprotein E dramatically reduces amyloid beta-peptide deposition. *Nat Genet* 1997; 17: 263–264.
84. Holtzman DM, Bales KR, Tenkova T, et al. Apolipoprotein E isoform-dependent amyloid deposition and neuritic degeneration in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci U S A* 2000; 97: 2892–2897.
85. Fagan AM, Watson M, Parsadanian M, et al. Human and murine ApoE markedly alters A beta metabolism before and after plaque formation in a mouse model of Alzheimer's disease. *Neurobiol Dis* 2002; 9: 305–318.
86. Refolo LM, Pappolla MA, LaFrancois J, et al. A cholesterol-lowering drug reduces beta-amyloid pathology in a transgenic mouse model of Alzheimer's disease. *Neurobiol Dis* 2001; 8: 890–899.
87. Wang J, Ho L, Zhao Z, et al. Moderate consumption of cabernet sauvignon attenuates Abeta neuropathology in a mouse model of Alzheimer's disease. *FASEB J* 2006; 20: 2313–2320.
88. Adlard PA, Perreau VM, Pop V, Cotman CW. Voluntary exercise decreases amyloid load in a transgenic model of Alzheimer's disease. *J Neurosci* 2005; 25: 4217–4221.
89. Patel NV, Gordon MN, Connor KE, et al. Caloric restriction attenuates Abeta-deposition in Alzheimer transgenic models. *Neurobiol Aging* 2005; 26: 995–1000.
90. Lazarov O, Robinson J, Tang YP, et al. Environmental enrichment reduces Abeta levels and amyloid deposition in transgenic mice. *Cell* 2005; 120: 701–713.
91. Solomon B, Koppel R, Hanan E, Katzav T. Monoclonal antibodies inhibit in vitro fibrillar aggregation of the Alzheimer beta-amyloid peptide. *Proc Natl Acad Sci U S A* 1996; 93: 452–455.
92. Schenk D, Barbour R, Dunn W, et al. Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse. *Nature* 1999; 400: 173–177.
93. Janus C, Pearson J, McLaurin J, et al. A beta peptide immunization reduces behavioural impairment and plaques in a model of Alzheimer's disease. *Nature* 2000; 408: 979–982.
94. Morgan D. Immunotherapy for Alzheimer's disease. *J Alzheimers Dis* 2006; 9(suppl): 425–432.
95. Bard F, Cannon C, Barbour R, et al. Peripherally administered antibodies against amyloid beta-peptide enter the central nervous system and reduce pathology in a mouse model of Alzheimer disease. *Nat Med* 2000; 6: 916–919.
96. DeMattos RB, Bales KR, Cummins DJ, et al. Peripheral anti-A beta antibody alters CNS and plasma A beta clearance and decreases brain A beta burden in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci U S A* 2001; 98: 8850–8855.
97. Brody DL, Holtzman DM. Active and passive immunotherapy for neurodegenerative disorders. *Annu Rev Neurosci* 2008; 31: 175–193.
98. Dodart JC, Bales KR, Gannon KS, et al. Immunization reverses memory deficits without reducing brain Abeta burden in Alzheimer's disease model. *Nat Neurosci* 2002; 5: 452–457.
99. Kotilinek LA, Bacskai B, Westerman M, et al. Reversible memory loss in a mouse transgenic model of Alzheimer's disease. *J Neurosci* 2002; 22: 6331–6335.
100. DeMattos RB, Bales KR, Cummins DJ, et al. Brain to plasma amyloid-beta efflux: a measure of brain amyloid burden in a mouse model of Alzheimer's disease. *Science* 2002; 295: 2264–2267.
101. Wisniewski T, Konietzko U. Amyloid-beta immunisation for Alzheimer's disease. *Lancet Neurol* 2008; 7: 805–811.
102. Orgogozo JM, Gilman S, Dartigues JF, et al. Subacute meningoencephalitis in a subset of patients with AD after Abeta42 immunization. *Neurology* 2003; 61: 46–54.
103. Hock C, Konietzko U, Streffer JR, et al. Antibodies against beta-amyloid slow cognitive decline in Alzheimer's disease. *Neuron* 2003; 38: 547–554.

Mount Sinai Journal of Medicine 77:82–102, 2010

Role of Vascular Risk Factors and Vascular Dysfunction in Alzheimer's Disease

Dara L. Dickstein, PhD,[1] Jessica Walsh, BA,[1] Hannah Brautigam, BSc,[1]
Steven D. Stockton Jr, BSc,[1] Samuel Gandy, MD, PhD,[2,3] and Patrick R. Hof, MD[1,4]

[1]Department of Neuroscience and Kastor Neurobiology of Aging Laboratories, Mount Sinai School of Medicine, New York, NY
[2]Department of Neurology, Mount Sinai School of Medicine, New York, NY
[3]Department of Psychiatry, Mount Sinai School of Medicine, New York, NY
[4]Department of Geriatrics and Palliative Care, Mount Sinai School of Medicine, New York, NY

OUTLINE

NEUROPATHOLOGY
  Multi-Infarct Dementia
  Lacunes
MIXED DEMENTIA
  Cerebral Amyloid Angiopathy
RISK FACTORS
  Apolipoprotein E Genotype
  Hypertension
  Hypotension
  Cholesterol
  Diabetes Mellitus
  Smoking
BLOOD-BRAIN BARRIER DYSFUNCTION
OXIDATIVE STRESS
CONCLUSION

ABSTRACT

Recent findings indicate that vascular risk factors and neurovascular dysfunction play integral roles in the pathogenesis of Alzheimer's disease. In addition to aging, the most common risk factors for Alzheimer's disease are apolipoprotein e4 allele, hypertension, hypotension, diabetes, and hypercholesterolemia. All of these can be characterized by vascular pathology attributed to conditions such as cerebral amyloid angiopathy and subsequent blood-brain barrier dysfunction. Many epidemiological, clinical, and pharmacotherapeutic studies have assessed the associations between such risk factors and Alzheimer's disease and have found positive associations between hypertension, hypotension, and diabetes mellitus. However, there are still many conflicting results from these population-based studies, and they should be interpreted carefully. Recognition of these factors and the mechanisms by which they contribute to Alzheimer's disease will be beneficial in the current treatment regimens for Alzheimer's disease and in the development of future therapies. Here we discuss vascular factors with respect to Alzheimer's disease and dementia and review the factors that give rise to vascular dysfunction and contribute to Alzheimer's disease. ***Mt Sinai J Med 77:82–102, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: Alzheimer's disease, apolipoprotein, blood-brain barrier, cholesterol, diabetes, hypertension, risk factors.

The prevalence of dementia, particularly Alzheimer's disease (AD), is increasing and it is one of the most important neurodegenerative disorders in the elderly. It is estimated that 5% to 10% of the elderly in the age range of 65 to 74 years are affected and that 25% to 50% of the elderly over the age of 85 years are affected. Of these cases, AD accounts for 50% to 60%, and the risk for AD doubles every 5 years after the age of 65 years.[1] It is projected that by 2040, 81.1 million people will be affected worldwide with dementia,[2] with frequencies for AD and vascular dementia (VaD) of 70% and 15%, respectively.[3] The neuropathological

Address Correspondence to:

Dara L. Dickstein
Department of Neuroscience
Mount Sinai School of Medicine
New York, NY
Email: dara.dickstein@mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20155

Abbreviations

$A\beta$	Amyloid beta protein
ACE	Angiotensin converting enzyme
AD	Alzheimer's disease
AGE	Advanced glycation end product
ApoE	Apolipoprotein E
APP	Amyloid precursor protein
AT1	Angiotensin II type 1
BACE	β-Site amyloid precursor protein cleavage enzyme
BBB	Blood-brain barrier
BP	Blood pressure
CAA	Cerebral amyloid angiopathy
CBF	Cerebral blood flow
CDR	Clinical dementia rating
CI	Confidence interval
CNS	Central nervous system
DBP	Diastolic blood pressure
DM	Diabetes mellitus
DSM-III	*Diagnostic and Statistical Manual of Mental Disorders*, 3rd edition
LRP	Low-density lipoprotein receptor–related protein
MMSE	Mini-Mental State Examination
MRC	Medical Research Council
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
NFT	Neurofibrillary tangle
NINCDS-ADRDA	National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association
NO	Nitric oxide
Nox2	gp91phox
OR	Odds ratio
PROGRESS	Perindopril Protection Against Recurrent Stroke Study
PUFA	Polyunsaturated fatty acid
RAGE	Receptor for advanced glycation end products
ROS	Reactive oxygen species
SBP	Systolic blood pressure
SCOPE	Study on Cognition and Prognosis in the Elderly
SHEP	Systolic Hypertension in the Elderly Program
SOD	Superoxide dismutase
Syst-Eur	Systolic Hypertension in Europe
TC	Total cholesterol
VaD	Vascular dementia

hallmarks of AD are intraneuronal protein aggregates of hyperphosphorylated tau protein [neurofibrillary tangles (NFTs)], aggregation of neurotoxic amyloid beta ($A\beta$) protein in the brain parenchyma and in blood vessel walls,[4] and neuronal degeneration and loss. Currently, one of the prevailing theories of AD pathophysiology is the amyloid cascade hypothesis, which implicates $A\beta$ as the key player in the formation of senile plaques and neuronal death.[4] However, recent evidence from epidemiological, clinical, pathological, and neuroimaging studies implicates neurovascular dysfunction as an integral part of AD and has given rise to the vascular hypothesis of AD. Moreover, data from these studies reveal a distinct association between vascular risk factors and AD. These include hypertension,[5] total cholesterol (TC), type II diabetes mellitus (DM),[6] hypotension, smoking,[7] and oxidative stress.[8] Furthermore, dysfunction of the endothelial cells that compose the blood-brain barrier (BBB) has also been demonstrated and correlates with AD severity.[9] The degree to which these factors contribute to AD may be influenced by genetic factors such as apolipoprotein E (ApoE), which has a role in both AD and vascular disease.

> *Recent evidence from epidemiological, clinical, pathological, and neuroimaging studies implicates neurovascular dysfunction as an integral part of Alzheimer's disease (AD).*

NEUROPATHOLOGY

Vascular pathology in the aging brain and AD includes ischemic infarcts, lacunes, cerebral hemorrhages, white matter lesions, BBB dysfunction, cerebral amyloid angiopathy (CAA), and microvascular degeneration.[10] These pathologies are commonly seen in various vascular diseases and can contribute to cognitive impairment by affecting neuronal networks involved in cognition, memory, behavior, and executive functioning.[11]

Multi-Infarct Dementia

It was first proposed that dementia occurs only in the presence of cerebral infarcts having a volume larger than 100 ml.[12] However, challenging this view, recent work has suggested that small infarcts and even multiple microscopic infarcts can also lead to dementia.[13–17] Although the accumulation of macro-infarcts, particularly in the hippocampus and the association neocortex, represents the primary form of VaD, much less is known regarding the contributions of microvascular pathology to cognitive decline leading to AD. Microvascular lesions indeed encompass a broad spectrum of pathologies, are highly heterogeneous, and may present with different patterns of clinical symptoms.[18]

Additionally, pathological lesions associated with AD, including an accumulation of Aβ and NFTs, may mask clinically relevant consequences due only to the microvascular pathology itself.[18–20]

Keeping these confounders in mind, many studies have attempted to elucidate the contributions of multiple micro-infarcts to the pathogenesis of dementia. The Nun study found that cerebrovascular disease determines the severity of AD and that patients with multiple lacunar infarcts had poorer cognitive function and a higher prevalence of dementia, regardless of the NFT load, in comparison with those without infarcts.[15] The Rotterdam Scan study found a similar association: patients with multiple brain infarcts had double the risk for developing dementia and a steeper decline in global cognitive function.[21] Recently, Kövari *et al.*[18] used postmortem human brain specimens taken from patients ranging in age from 63 to 100 years and found that the presence of micro-infarcts as well as focal cortical gliosis had a significant correlation with the extent of cerebral microvascular pathology and cognitive function.[18]

> *The Nun study found that cerebrovascular disease determines the severity of AD and that patients with multiple lacunar infarcts had poorer cognitive function and a higher prevalence of dementia, regardless of the neurofibrillary tangle load, in comparison with those without infarcts.*

Lacunes

Lacunes, found in the periventricular white matter and in subcortical structures such as the basal ganglia, internal capsule, thalamus, pons, corona radiata, and centrum semiovale, are small ischemic infarcts surrounded by reactive gliosis and macrophages that have a diameter of less than 15 mm.[22–24] Such lesions may present clinically as severe dysexecutive syndrome accompanied by deficits in memory, frontal lobe deficits, slowing of motor function, changes in personality, impairment in activities of daily living, parkinsonian features, and other problems associated with deterioration of working memory (such as deficits in the planning, organization, and sequencing of events).[25]

Studies examining the clinical significance of lacunes as well their effects on cognition have yielded largely mixed results. Although as many as 23% of individuals older than 65 years of age may present with lacunes, up to 25% of these lacunes appear to be clinically silent: they have little or no observable effect on cognitive functioning.[22] However, when lacunes are observed in patients presenting with concomitant AD-related alterations, the presence of deep white matter, basal ganglia, and thalamic lacunes significantly affects cognitive function.[15] In a comparative autopsy study, it was found that there was a higher incidence of cerebrovascular lesions in AD patients in comparison with age-matched controls (48.0% versus 32.8%, $P < 0.01$) with a high incidence of minor to moderate lacunar lesions.[26] As the emergence of AD-related pathology can complicate the relationship between vascular lesions and dementia, Gold *et al.*[23] controlled for the presence of NFTs by including only those cases that met the criteria for Braak stages I and II and thus ensured that AD-related pathologies would be less likely to mask the contributions of lacunar and microvascular pathologies to clinical dementia rating (CDR) scores. The results showed that cognitive status was significantly predicted by the presence of lacunes in the thalamus and basal ganglia (5% and 6%, respectively). Interestingly, this study also documented significant interactions between Aβ staging (7% of CDR variability), age (3% of CDR variability), and various microvascular pathologies (23% of CDR variability) and variability in CDR scores.[22,24]

> *In a comparative autopsy study, it was found that there was a higher incidence of cerebrovascular lesions in AD patients in comparison with age-matched controls (48.0% versus 32.8%, $P < 0.01$) with a high incidence of minor to moderate lacunar lesions.*

MIXED DEMENTIA

Across the lifespan, normal brain aging includes the development of various AD-related pathologies as well as the progressive emergence of lesions associated with a range of vascular pathologies. For the majority of the aging population, the cognitive impact of this combination of pathologies remains essentially silent.[27] The emerging concept of mixed dementia refers to a broad spectrum of conditions in which cognitive declines may be attributed to the presence of both AD and vasculature-related alterations.[11,26] At present, research efforts have been made more difficult as a result of the broad application of this diagnosis; patients presenting with every combination of pathologies from severe AD lesions but sparse vascular pathology to severe vascular pathology with few AD lesions are eligible for the diagnosis of mixed dementia.[28] Additionally, a range of structural alterations of microvessels have been documented in both normal aging and AD. Normal aging has been associated with both microvascular lesions and the presence of amyloid deposition and NFTs in nondemented patients with hypertension, whereas patients with AD exhibit severe vascular modifications, including atrophic or string vessels, glomerular loop formation, twisted vessels, and fragmentation of the microvasculature.[28,29] Hampering research efforts still further is the lack of well-defined threshold values for both AD and vascular lesions that consistently predict cognitive status and dementia in patients as well as the diffuse nature of the pathological lesions under study. In an effort to develop better thresholds for the diagnosis of mixed dementia, Gold and colleagues[27] again used systematic semiquantitative methods both to control for the confounding effects of age and to more reliably identify cutoff values that consistently identify individuals as demented or nondemented on the basis of their burden of AD and vascular pathology.

The emerging concept of mixed dementia refers to a broad spectrum of conditions in which cognitive declines may be attributed to the presence of both AD and vasculature-related alterations.

These authors assessed the relationship between CDR scores of a randomized sample of patients and the presence of the aforementioned previously identified neuropathological features. Following the univariate analysis, these data[24] closely matched previous findings[23,27] and suggested that Braak NFT staging, Aβ deposition staging, cortical micro-infarct scores, the degree of lacunar pathology in the thalamus and basal ganglia, and age were all significantly related to CDR scores. When the data were further parsed with a multiple regression model, the effect of age was lost, and up to 48.9% of the variability in the CDR scores could be explained by the variations in NFT staging, Aβ staging, cortical micro-infarcts, and lacunes alone. Among the most important findings of this work, however, was the identification of neuropathological thresholds that the authors were able to use to classify accurately up to 90% of their sample as demented or nondemented. Moreover, the derivation model employed by Gold *et al.*[27] resulted in high sensitivity and specificity, a strong positive predictive value, and a high correct rate of classification.

Cerebral Amyloid Angiopathy

CAA is defined as the deposition of Aβ peptide within the walls of the leptomeninges and parenchymal arteries, arterioles, and capillaries with a concomitant thickening of arteriole walls and formation of micro-aneurysms. In addition, CAA has been associated with the degeneration of smooth muscle cells, ischemic white matter damage, fibrinoid necrosis, and dementia (reviewed by Jellinger[11]). The majority of CAA is spontaneous and its incidence increases with age to almost 100% past the age of 80. In AD, CAA can range from 70% to 97.6%.[30] The origin of the Aβ deposited in blood vessels has not been clearly elucidated. One possibility, the drainage hypothesis, suggests that neurons are the main source of vascular amyloid because neurons are the main source of amyloid precursor protein (APP) and Aβ in the brain.[31] Normally, Aβ produced in neurons is transported across the BBB. However, if there is a deficiency in the transport mechanism of Aβ, Aβ can build up in the vessel walls. It has been shown that the effect of widespread Aβ deposition is the degeneration and death of endothelial cells and the obliteration of the capillary lumen.[32] Moreover, ultrastructural studies indicate that approximately 32% of fibrillar amyloid plaques are in contact with 1 or more cerebral capillaries[33] and that 77% of plaques in Tg2576 mice and 91% of human plaques are in direct contact with capillaries.[33] Another theory is that Aβ can be systemic and originate in the circulating bloodstream. Finally, it has also been suggested that smooth muscle cells within the vessel walls or pericytes produce the Aβ.[34] Regardless of its origins, the Aβ associated with CAA, causing microscopic bleeding throughout the neocortex and its associated lobar white matter, has been shown to have a pathogenic role in dementia (reviewed by Jellinger[11]). Many of the mouse models for AD, harboring many different APP mutations associated with AD (ie, Swedish, Dutch, Iowa, and London), have also been shown to exhibit CAA. In addition to the accumulation of parenchymal Aβ, many of these models have Aβ deposits within the vessel walls of the leptomeninges and the neocortical, hippocampal, and thalamic vessel walls.[35–44] Therefore, it would appear that CAA may be an integral cerebrovascular dysfunction in the diagnosis of AD and dementia.

RISK FACTORS

AD is a complex, multifactorial neurodegenerative disorder likely resulting from the contribution of complex gene-gene and gene-environmental interactions. There are many underlying risk factors that contribute to vascular disease and AD, including apolipoprotein genotype, hypertension, hypotension, cholesterol levels, DM, and smoking. These, in addition to environmental risk factors (ie, brain injury, metals, education levels, dietary factors, and smoking), have been studied and implicated as possible risk factors for AD.

There are many underlying risk factors that contribute to vascular disease and AD, including apolipoprotein genotype, hypertension, hypotension, cholesterol levels, diabetes mellitus, and smoking.

Apolipoprotein E Genotype

Although there are many environmental risk factors that contribute to dementia in AD, genetic risk factors, such as ApoE, play a central role in its pathophysiology. In addition to its association with sporadic AD, ApoE is also related to vascular disease. ApoE is a plasma cholesterol transport molecule found on chromosome 19q13.2, and it occurs in 3 common alleles ($\epsilon 2$, $\epsilon 3$, and $\epsilon 4$).[45,46] It is a key constituent in very low density lipoproteins and is vital in the transport of cholesterol and other lipids throughout the brain.[47] ApoE in the central nervous system (CNS) is expressed primarily in astrocytes but can also be found in microglia and neurons. It is thought to act as a neurotrophic factor in growth and repair during CNS development and injury and is regulated by endocytosis with low-density lipoprotein receptor–related protein 1 (LRP1).[48]

Many epidemiological studies have found an association between ApoE $\epsilon 4$ and AD.[49] ApoE $\epsilon 4$ is known to play a major role in AD and dementia and is associated with a younger age of onset in a dose-dependent manner.[50] Those homozygous for the ApoE $\epsilon 4$ allele have a 12-fold increase in the risk for AD.[49] However, ApoE $\epsilon 4$ is neither necessary nor sufficient for the development of AD. ApoE $\epsilon 4$ is also known to be associated with coronary heart disease[51] and has been implicated as a risk factor for atherosclerosis[52] and stroke.[53,54] ApoE $\epsilon 4$ has also been shown to have a strong association with CAA, and it has been suggested that the contribution of CAA to AD is largely dependent on $\epsilon 4$.[55] Data from the Honolulu-Asia study and the Rotterdam study indicate that patients who are $\epsilon 4$ carriers have more pronounced vascular risk factors than noncarriers.[56] Moreover, patients who are $\epsilon 4$ carriers and have atherosclerosis are at increased risk of cognitive decline.[57] It has also been shown in a different population cohort that patients with mild AD who are $\epsilon 4$ carriers have a faster rate of cognitive decline.[58] Studies in transgenic mice have also demonstrated that ApoE $\epsilon 4$ promotes the formation of CAA.[44,59]

The mechanisms by which ApoE contributes to AD are not completely understood. Many studies have demonstrated that ApoE has isoform-specific capabilities ($\epsilon 2 > \epsilon 3 > \epsilon 4$) for acting as a chaperone molecule for Aβ and influences Aβ metabolism, deposition, toxicity, fibril formation, and clearance from the brain.[60–64] In addition, ApoE could also mediate tau hyperphosphorylation and modulate the distribution and metabolism of cholesterol in neuronal membranes in an isoform-dependent manner.[46] This is supported by evidence showing that increased plasma cholesterol concentrations correlate with increased Aβ accumulation in the brains of humans and transgenic mice and may be a result of serum ApoE concentrations.[65]

Hypertension

The relation between high blood pressure, cognitive function, and dementia has been the subject of numerous epidemiological and clinical studies that have generated a rather mixed outcome. Hypertension, currently defined as a systolic blood pressure (SBP) above 140 mm Hg and/or a diastolic blood pressure (DBP) above 90 mm Hg,[66,67] is a risk factor for many disorders, including AD, stroke, atherosclerosis, myocardial infarction, and cardiovascular disease.[68] Hypertension is estimated to affect

DOI:10.1002/MSJ

25% of the general population with a 50% prevalence in people over 70 years of age.[68] Epidemiologically, it has been shown that hypertension precedes dementia onset by approximately 30 years; however, this relationship is complex and does not follow a linear progression.[69] Midlife hypertension is particularly associated with an increased risk of developing both AD and VaD, whereas elevated blood pressure late in life does not appear to have the same associated risk. By itself, hypertension has been shown to be an independent risk factor for AD, but it is also associated with other diseases such as cardiovascular disease and stroke, which are known to be important factors leading to the onset of dementia.[70] Neuropathological and imaging investigations have demonstrated that individuals with high blood pressure often have large areas of white matter hyperintensity (manifested histopathologically as demyelination, arteriolosclerosis, gliosis, and tissue degeneration), ventricular enlargement, and silent infarcts, all of which can lead to cognitive dysfunction and dementia.[1]

Many population-based studies, both cross-sectional and longitudinal, have evaluated the link between hypertension and memory impairment and have generated conflicting data. Several longitudinal studies have confirmed that hypertension or elevated blood pressure, occurring in middle age or late in life, plays an important role in the development of cognitive dysfunction and is associated with an increase in the risk for AD and dementia.[71–74] However, other studies have shown that treatment with antihypertensive drugs has no significant effect on AD. One of the first studies that provided evidence relating blood pressure and cognitive decline was the Framingham study. This study concluded that elevated blood pressure was associated with modest impairment of cognitive function.[75,76] Following this study, several randomized placebo-controlled clinical trials evaluated the effect of antihypertensive drugs on dementia and AD (summarized in Table 1). The Rotterdam study, the Kungsholmen study, the Honolulu-Asia Aging Study, and the Epidemiology of Vascular Aging Study supported the results observed in the Framingham study.

In addition to the epidemiological studies associating hypertension with the incidence of AD, many studies have found associations between hypertension and AD brain pathology. Data from a brain imaging study found cross-sectional associations between SBP and medial temporal lobe atrophy in patients with AD. Hippocampal atrophy has also been reported in both the Honolulu-Asia study and the Rotterdam study in patients not treated for hypertension. Furthermore, the Honolulu-Asia study also reported an association between midlife SBP, lower brain weight, and an abundance of amyloid plaques in the hippocampus and neocortex. Furthermore,

Table 1. *Summary of Hypertensive Therapy Studies in AD.*

Reference	Study Setting	Participants and Follow-Up	Treatment	Main Result
232	SHEP	4736 people, ~72 years at baseline, followed up to 4.5 years	Diuretic ± β-blocker ± hypertensive drug or placebo	No significant effect of treatment on AD risk
233, 234	MRC trial	2584 people, ~69 years at baseline, followed up to 3.9 years	Diuretic or β-blocker or placebo	No significant effect of treatment on AD risk
235, 236	SCOPE	4964 people, ~76 years at baseline, followed up to 4.5 years	ATI receptor agonist or placebo (candesartan ± hydrochlorothiazide versus placebo)	Significantly less cognitive decline in the treated group
237	PROGRESS	6105 people, ~64 years at baseline, followed up to 4.5 years	ACE inhibitor ± diuretic or placebo (perindopril ± indapamide versus placebo)	Decreased the rates of dementia in actively treated patients
238–240	Syst-Eur study	2418 people, ~68 years at baseline, followed up to 2 years	Calcium channel blocker ± ACE inhibitor ± diuretic or placebo (nitrendipine ± enalapril ± hydrochlorothiazide versus placebo	Decreased the risk of dementia by 55%

NOTE: This table is a summary of randomized, double-blind, placebo-controlled studies describing the association between hypertension and antihypertensive therapy in relation to AD and has been adapted from Takeda *et al.*[1] and Poon.[241]

Abbreviations: ACE, angiotensin converting enzyme; AD, Alzheimer's disease; ATI, angiotensin II type 1; MRC, Medical Research Council; PROGRESS, Perindopril Protection Against Recurrent Stroke Study; SCOPE, Study on Cognition and Prognosis in the Elderly; SHEP, Systolic Hypertension in the Elderly Program; Syst-Eur, Systolic Hypertension in Europe.

patients with a DBP > 95 mm Hg exhibited higher numbers of NFTs in the hippocampus.[77] It is thought that the onset of these pathologies occurs prior to the onset of dementia as plaques and NFTs were also present in nondemented, middle-aged individuals with hypertension.[77] A few theories about how hypertension can contribute to AD and dementia have been proposed. First, it is thought that hypertension causes vascular alterations that then lead to lacunar or cortical infarcts and leukoaraiosis and ultimately cognitive decline. Second, it has been suggested that hypertension leads to cardiovascular disease, which gives rise to AD.[15] Third, hypertension can have adverse effects on neuronal health and increase the production of Aβ and can thereby lead to neuronal dysfunction, synapse and neuronal loss, and dementia.[78] Altogether, it seems that hypertension, aging, and cerebrovascular risk factors act synergistically to cause vascular degeneration, oxidative stress, mitochondrial dysfunction, neuronal degeneration, and AD.

Hypotension

Although it has been shown that increased blood pressure is a strong risk factor for AD and VaD, a decrease in blood pressure can also have adverse effects on cognition in old age.[79–81] Hypotension, defined as having a DBP ≤ 70 mm Hg, is usually associated with increased mortality.[82] However, recent studies have shown that hypotension is a key factor in conditions such as diabetes and psychosomatic distress,[83,84] which all can be considered risk factors for AD. The mechanism by which low blood pressure can lead to AD is speculative, but it is thought either to be a result of the dementia process or to accelerate and predispose to cognitive decline.[85]

Many population-based prospective studies have demonstrated an increased prevalence of AD and dementia in persons with low blood pressure (Table 2).[79–81] Reports from the Honolulu-Asia study found that in a cohort of Japanese-American men, low DBP was associated with an increased risk of developing AD in later life. This risk was particularly significant in subjects that were not treated with antihypertensive drugs.[73] Further longitudinal studies of cohorts from Sweden (the Kungsholmen study) and Boston (the East Boston study) found that low DBP (<70 mm Hg) tended to increase the relative risk for AD and dementia.[81,86] In another population-based cohort, Zhu *et al.*[80] examined 924 persons who were 75 years old or older and found that there was also a correlation between systolic pressure reduction and cognitive decline in women.[80] Results from the Kungsholmen project and the Chicago Health and Aging Project corroborated these data and showed that people with an SBP ≤ 140 mm Hg and a DBP < 70 mm Hg or an SBP ≤ 130 mm Hg and a DBP < 70 mm Hg had significantly higher risks of

Table 2. *Studies Describing the Association Between Hypotension and Dementia.*

Reference	Study Setting	Participants and Follow-Up	Outcomes	Main Results
89	Gothenburg H-70 and Rotterdam studies	382 people, 70 years, followed up to 15 years	Dementia, AD; DSM-III, NINCDS-ADRDA	BP was inversely related to the risk of dementia in patients taking antihypertensive drugs.
81	East Boston study	378 people, ≥65 years, followed up to 3 years	AD; NINCDS-ADRDA	SBP ≥ 160 mm Hg versus 130–139 mm Hg (OR = 0.3, 95% CI = 0.1–0.9); DBP < 70 mm Hg versus 80–89 mm Hg (OR = 1.8, 95% CI = 0.3–4.3)
79	Kungsholmen study	1642 people, ≥75 years	Dementia, AD; DSM-III	People with DBP < 70 mmHg or SBP < 140 mm Hg had a higher risk of dementia and AD.
88	Chicago Health and Aging Project	709 people, ≥65 years	AD; NINCDS-ADRDA	People with DBP < 70 mmHg or SBP < 130 mm Hg had a higher risk of dementia and AD.
242	OCTO-Twin study	599 people, ≥80 years at baseline, mean follow-up of 4 years	Dementia; DSM-III, MMSE	Lower SBP and DBP were associated with cognitive decline.

Abbreviations: AD, Alzheimer's disease; BP, blood pressure; CI, confidence interval; DBP, diastolic blood pressure; DSM-III, *Diagnostic and Statistical Manual of Mental Disorders*, 3rd edition; MMSE, Mini-Mental State Examination; NINCDS-ADRDA, National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association; OR, odds ratio; SBP, systolic blood pressure.

dementia and AD, respectively.[87,88] Moreover, the risk of AD associated with low blood pressure was particularly pronounced in antihypertensive drugs users.[86] More recently, Ruitenberg *et al.*[89] observed that there was a higher prevalence of decreased blood pressure, independent of age and sex, in demented patients in comparison with nondemented patients at follow-up.

How hypotension is associated with the risk for AD remains unclear. It is possible that low blood pressure occurs as a result of brain pathology. Pathological changes, such as the development of Aβ plaques, can lead to a reduction of arterial pressure, which in turn may produce hypoxic-ischemic changes that would act synergistically with existing pathology to exacerbate the degree of dementia.[72,79,90]

Alternatively, low blood pressure may predispose people to the risk of AD and dementia and act as an early correlate of the dementia process.[81,86,89] Aging of the vasculature results in changes in the structural and mechanical properties of the arterial walls that ultimately lead to a dampening of the autoregulatory capabilities of cerebral arteries rendering the brain more vulnerable to ischemia and hypoperfusion.[91] Evidence shows that a decrease in cerebral blood flow (CBF) precedes the neuropathology observed in AD, and it continues to decline during the course of the disease.[92,93]

Cholesterol

The specific mechanisms underlying the correlation between cholesterol metabolism and AD are controversial.[47,94] It has been suggested that cholesterol plays an essential role in regulating the enzyme activity that is involved in the production of Aβ protein and the metabolism of APP.[95] In AD, the cleavage of APP occurs within the hydrophobic lipid bilayer and is catalyzed by the activity of the α-secretase, β-secretase, and γ-secretase enzymes. As such, disturbances in the levels of cellular cholesterol, which cause disorganization in the structure of the lipid bilayer, could alter the processing of APP by α-secretase by shifting the proximity of the secretase cleavage sites to the intramembrane domain of APP.[96] The activity of β-secretase [or β-site amyloid precursor protein cleavage enzyme (BACE)] and γ-secretase appears to be dependent on the composition of lipid rafts in the membrane.[97–99] Many in vitro studies have demonstrated that high levels of cholesterol affect α-secretase and BACE activity and result in a decrease in soluble APP levels and an increase in Aβ_{1-40} and Aβ_{1-42}. Conversely, cholesterol depletion can promote α-secretase activity and the production of soluble APP while decreasing the production of Aβ_{1-40} and Aβ_{1-42}.[97,100–103] The impact of cholesterol on γ-secretase function is still unresolved. It has been shown that γ-secretase is dependent on lipid rafts but is not cholesterol-dependent,[103,104] whereas others have shown that cholesterol can indeed modulate enzyme activity.[98,105,106] It has been proposed that high levels of cholesterol alter the plasma-membrane composition and appear to impede membrane fluidity and thereby prevent the interaction of α-secretase with APP and preclude the production of soluble APP while increasing the number of lipid rafts in the membrane and thereby facilitating the interaction between APP and BACE and the generation of Aβ.[100] This shift of APP within the membrane and the increase or decrease in lipid rafts in the membrane can lead to neuronal degeneration because soluble APP is neuroprotective and has been shown to act as a trophic factor, reduce intracellular calcium concentrations, and protect against hypoglycemic damage and glutamate toxicity,[97,107] whereas Aβ is known to be neurotoxic.

In vivo studies have also demonstrated the effect of cholesterol on the generation of Aβ. Staining of hippocampal and frontal cortex paraffin sections showed that rabbits given high cholesterol diets had both high-intensity staining and increased levels of brain Aβ. When the animals were returned to a normal diet, the intensity of staining was significantly reduced, and the levels of brain Aβ returned to normal.[108,109] Recent findings have suggested that a partial loss of γ-secretase function and an accumulation of γ-secretase substrates impair endocytosis of lipoproteins.[110] Refolo *et al.*[111] showed that diets high in cholesterol, administered to rabbits and AD mouse models, increased Aβ levels.

Several epidemiological studies analyzing the possible connections between AD and cholesterol levels have conflicting results, but the majority stress an association between high plasma cholesterol in midlife and increased susceptibility to sporadic AD.[47,112–114] Interestingly, there was no significant correlation found when cholesterol levels were analyzed in patients later in life.[114,115] In an investigation by Kuusisto *et al.*,[116] 980 participants between the ages of 69 and 78 years had a positive association between low serum TC and AD. In contrast, another analysis of 1449 elderly participants suggested that there was an inverse relationship between decreased TC and the incidence of AD.[117] Furthermore, a study demonstrated that in a group of individuals lacking an ApoE ϵ4 allele, high TC was associated with AD.[118] Other studies looking at the use of 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors (statins) have shown that the use of statins decreases the incidence and prevalence of AD.[119–121] Statins inhibit 3-hydroxy-3-methylglutaryl coenzyme A reductase, the rate-limiting enzyme in the cholesterol biosynthetic pathway, in addition to affecting intracellular cholesterol distribution. Statins have been shown to alter the expression of genes involved in cell growth and signaling with a neuroprotective effect.[122] Initial studies suggested that the probable risk of AD development in patients who were taking statins was lower than that in those patients taking other medications.[120] However, other studies have shown that statins may not have a protective effect against dementia. The Prospective Study of Pravastatin in the Elderly at Risk did not find statins to have a significant effect on the cognitive functions of patients ranging from 70 to 82 years of age.[123] The Heart Protection Study, which included 20,536 patients who received either 40 mg of daily statins or a placebo, showed no statistical difference in cognitive decline. However, dementia was determined on the basis of phone interviews, which may have lower reliability in assessing dementia severity.[123]

The relationship between AD and cholesterol varies; it depends on when in life the levels of cholesterol are measured. High cholesterol in midlife seems to be a high risk factor for developing AD; however, individuals who have high cholesterol in late life do not seem to present with an increased risk of dementia. Statins have been an interesting avenue of research as they are the first line of treatment for hypercholesterolemia and because of evidence that they may prevent neuronal death.[95] Although the aforementioned studies show that cellular cholesterol levels modulate APP, the exact mechanisms through which this modulation occurs remain elusive.

Diabetes Mellitus

DM is a metabolic disorder that is common in more than 10% of the elderly population in the United States[95] and is associated with changes in mental cognition and flexibility.[124] Type 1 DM is characterized by a deficit in the production of insulin by the pancreatic β cells. A review of 33 longitudinal studies suggested that patients with type 1 DM show an increase in cognitive dysfunction and a decrease in the speed of mental processing.[125] Type 2 DM is characterized by resistance to the effects of insulin. Longitudinal investigations have confirmed that there is a relationship between type 2 DM and a faster rate in the decline of cognition.[126]

Early studies suggested that there is no apparent correlation between type 2 DM and AD.[127–129] In fact, these studies, showing that patients with AD have a low rate of diabetes, suggest that there may not in fact be a link between diabetes and AD.[127] In contrast, others have reported that up to 80% of patients with AD also exhibit type 2 DM.[6,130–132] Longitudinal population-based studies that assessed both diabetes and dementia in late life found that the incidence of dementia was increased by 50% to 100% in diabetic patients.[124] Furthermore, prospective and cross-sectional analyses have proposed an association between type 2 DM and an increased risk of AD, particularly in poststroke patients; they suggest that diabetes may accelerate the onset of AD, rather than increasing the long-term risk.[133]

Three pathophysiological mechanisms have been proposed to explain the association between DM and dementia.[124] First, diabetic individuals have an increased risk of developing dementia through ischemic cerebrovascular disease. Type 2 DM, which is most prevalent in elderly individuals, can develop a cluster of risk factors such as insulin resistance, obesity, and hypertension, which can constitute a metabolic syndrome.[116,134,135] This combination of risk factors has been established as a predictor of cerebrovascular disease and dementia.[136]

Second, it has been proposed that hyperglycemia has toxic effects on neurons, which can in turn lead to functional or cellular brain deficits through oxidative stress and the accumulation of glycation end products.[137,138] Advanced glycation end products (AGEs) are sugar-derived substances formed by a nonenzymatic reaction between reducing sugars and free amino groups of proteins, nucleic acids, and lipids. They are normally produced in the body; however, their formation is greatly increased in individuals with diabetes because of the increased glucose availability.[139] Data have suggested that the primary event initiating both intracellular and extracellular AGEs is intracellular hyperglycemia.[140] AGEs can come from intracellular auto-oxidation, which forms reactive dicarbonyls such as glyoxal, methylglyoxal, and 3-deoxyglucosone.[141,142] There are 3 different mechanisms through which target cells can be damaged by the production of intracellular AGE precursors. First, intracellular proteins have been modified covalently by dicarbonyl AGE precursors, which alter several cellular functions. For example, the basic fibroblast growth factor found in endothelial cells is one of the predominant AGE-modified proteins.[143] Proteins that play a role in the endocytosis of macromolecules are also modified by AGEs, as overexpression of the methylglyoxal-detoxifying enzyme glyoxalase I prevents the increase in endocytosis caused by hyperglycemia.[144] Also, AGE formation causes abnormal interactions of modified extracellular matrix proteins with other matrix proteins and integrins,[139] and this results in decreased elasticity of vessels in diabetic rats, even when vascular tone is abolished.[145] Moreover, plasma proteins modified by AGE precursors produce ligands that bind to AGE receptors on endothelial cells.[146] Such binding to the AGE receptor induces the activation of a transcription factor known as nuclear factor kappa B, which causes pathological changes in the expression of several genes, such as the expression of proinflammatory molecules by endothelial cells.[147,148] Furthermore, endothelial AGE receptors that have bound ligands partially result in hyperpermeability of the capillary wall, which is caused by diabetes.[149] Recent evidence has, therefore, suggested that complications caused by diabetes are fueled by glucose and oxidative, proinflammatory AGEs.[150]

Lastly, resistance to insulin is associated with hyperinsulinemia, which is a risk factor for dementia.[116,151] Because insulin has vasoactive effects, this association is at least partly mediated through vascular disease.[152] Furthermore, observations have suggested that insulin may have a direct effect on the brain.[153–155] Insulin is actively transported across the BBB and may also be locally produced in the brain.[156,157] Insulin receptors are widely present in many regions of the brain, such as the granule cell layer of the olfactory bulb, the cerebellar cortex, and the hippocampal formation.[158] Insulin in the brain is also a modulator of energy homeostasis and intake of food.[159] In fact, patients with AD may have impaired activation of insulin receptors in the brain, and this suggests that AD could in fact be considered an insulin-resistant brain state.[160,161] It appears that the metabolism and removal of Aβ are directly affected by insulin.[162] Aβ breakdown through the insulin-degrading enzyme is inhibited by insulin.[163] Insulin stimulates Aβ intracellular trafficking in neuronal cultures and thereby directly increases the secretion of Aβ and decreases the intracellular levels of Aβ peptides.[164] Investigations in Tg2576 mice found that diet-induced insulin resistance can increase AD-type amyloidosis in the brain through an impairment of insulin receptor signaling resulting in an increase in γ-secretase activity.[165] These mechanisms may provide a substrate for the apparent association between diabetes and AD. Furthermore, they suggest that vascular disease, changes in glucose blood concentration, and amyloid metabolism are important factors in understanding how diabetes affects brain function, particularly in individuals with AD.

Smoking

The relationship between smoking and neurodegenerative diseases is controversial. Previous predominantly case-controlled studies have proposed that the nicotine in cigarette smoke is inversely correlated with AD and cognitive decline.[166–168] In contrast, it has also been established that smoking causes many deleterious effects through vascular mechanisms, which result in atherosclerosis and thrombosis and increase the risk of AD.[169,170] Furthermore, although some studies have reported weak or negative results,[171–173] others have suggested that former smokers are at a low risk of AD in comparison with nonsmokers[174,175] or are at a higher risk.[7,176]

The neuroprotective effects of smoking have been evaluated by many groups, and it has been found that cigarette smokers are 50% less likely to develop AD than age-matched and gender-matched nonsmokers (reviewed by Fratiglioni and Wang[177]). Prospective cohort studies have shown that there is an increased risk[7,175,178] or an unchanged risk[172,173,179] of AD in smokers. When comparing the risk of current smokers and former smokers to the risk of individuals who never smoked, the Chicago Health and Aging Project study found that smokers were 3 times more likely to develop AD than nonsmokers and that former smokers had no difference in risk in comparison with nonsmokers.[180] Studies examining the relationship between smoking and ApoE $\epsilon 4$ found that individuals who were smokers and lacked the ApoE $\epsilon 4$ allele had a greater relative risk of developing AD, whereas those who had ApoE $\epsilon 4$ had an insubstantial relative risk.[181]

Various possible mechanisms associating smoking history with the onset of AD have been proposed. First, it has been established that exposure to tobacco can lead to the development of atherosclerosis, which in turn leads to an increased risk of ischemic stroke[169,170] and hence dementia. Second, nicotine has been shown to modulate the neurotoxic effects of Aβ, can exert potent neuroprotective effects, and may confer resistance to AD. The presence of nicotine from cigarettes causes an up-regulation and activation of nicotinic acetylcholine receptors, which in turn protect against Aβ cytotoxicity.[168,182–184] Cholinergic deficits, characterized by reduced levels of acetylcholine nicotinic receptors, are found in AD.[185] Furthermore, Poirier *et al.*[186] found that patients with AD who have an ApoE4 allele have fewer nicotinic receptor binding sites and a depletion of choline. Third, nicotine can also have an effect on oxidative stress. In vitro and in vivo studies have demonstrated that nicotine can act as a scavenger of oxygen free radicals and efficiently scavenge superoxide and hydroxyl radicals in the brain.[187] Other possible roles of nicotine include its ability to inhibit arachidonic acid–induced apoptosis cascades,[188,189] Aβ-induced apoptosis,[187] *N*-methyl-d-aspartate receptor–mediated calcium-dependent excitotoxicity,[190] and Aβ fibril formation.[191] Transgenic mouse studies have also yielded conflicting results. Studies in transgenic mouse models have shown contradictory results: chronic nicotine administration has been shown to be effective in reducing total Aβ levels in the brain and improving cognition[35,192–195] or to have no effect at all.[196,197] Consequently, more mechanistic studies are needed to resolve whether smoking is protective or a detrimental risk factor for AD.

BLOOD-BRAIN BARRIER DYSFUNCTION

Another possible hypothesis that can account for the pathogenesis of AD is the impairment of the BBB.[198] It is known that cerebral blood vessels undergo profound changes with aging. These changes continue and are exacerbated in AD and have led to intensive research into the properties of the BBB. The BBB, found in all vertebrates, prevents the free diffusion of circulating molecules, leukocytes, and red blood cells into the brain interstitial space and is an essential regulator of the neuronal and glial cell environment. The barrier is formed by the presence of epithelial-like, high-resistance tight junctions that fuse brain capillary endothelial cells together into a continuous cellular layer separating the blood and brain.[199] The disruption of tight junctions that are found in endothelial cells results in altered transport of molecules between the blood and brain and the brain and blood, aberrant angiogenesis, vessel regression, brain hypoperfusion, and inflammatory responses, and it can have detrimental effects on synaptic plasticity and neuronal survival. Indeed, reduced microvascular density, increased fragmentation of vessels, increased thickening of basement membranes, increased vessel diameter, and a reduced number of endothelial mitochondria have all been described in AD (reviewed by Zlokovic[200]).

Disturbances in the BBB have been associated with stroke,[198] cerebrovascular ischemia,[198] hypertension,[201] and mutations in the ApoE gene.[202] The integrity of the BBB in aging and in AD is an area of contention and conflicting results. Given that the majority of AD cases are sporadic and do not have evidence of genetic mutation in APP and thus overproduction of Aβ, it is thought that the accumulation of Aβ in AD brains and on blood vessels is due to a deficiency in amyloid clearance from the brain.[203] This can occur because of either deficient efflux of amyloid from the brain or faulty degradation in the CNS that leads to the accumulation of neurotoxic amyloid in the brain (reviewed by Zlokovic[203]). There are 2 main receptors that are responsible for amyloid influx and efflux across the BBB. The receptor for advanced glycation end products (RAGE), responsible for the influx of amyloid across the BBB, is a multiligand receptor in the immunoglobulin superfamily found in neurons, microglia, and cerebral endothelial cells. It binds many ligands such as Aβ, the S100/calgranulin family of proinflammatory cytokine-like mediators, and the high-mobility group of DNA binding protein amphoterin.[204,205] It is normally expressed at low levels in the brain and is dependent on the presence of its ligands. In AD, RAGE expression is increased several-fold in affected vessels, neurons, and glia.[204,206] Studies in transgenic mice overexpressing mutated APP as well as RAGE have shown that RAGE is a cofactor for Aβ-induced neuronal perturbation in AD.[207,208] RAGE/Aβ complexes have been shown to decrease CBF[207] and initiate oxidative stress by activating microglia.[204] In RAGE/APP mice, early abnormalities in spatial learning, accompanied by altered activation of markers of synaptic plasticity and exaggerated neuropathology, were found. These changes were observed approximately 3 to 4 months earlier in comparison with APP mice.[208] Moreover, in APP transgenic mice bearing a dominant-negative RAGE construct, there was preservation of spatial learning/memory as well as diminished neuropathological changes.[208]

> *Given that the majority of AD cases are sporadic and do not have evidence of genetic mutation in amyloid precursor protein and thus overproduction of amyloid beta, it is thought that the accumulation of amyloid beta in AD brains and on blood vessels is due to a deficiency in amyloid clearance from the brain.*

LRP, responsible for the efflux of Aβ across the BBB, is a member of the low-density lipoprotein receptor family and is a multifunctional scavenger and signaling receptor. Its ligands include biomolecules such as ApoE, APP, Aβ,

α_2-macroglobulin, tissue plasminogen activator, and lactoferrin.[203] LRP is expressed in brain capillary endothelium and exhibits reduced expression during normal aging and AD.[206,209] Studies in LRP-deficient mice and AD transgenic mice have demonstrated a significant increase in the cerebral amyloid load and parenchymal plaques in comparison with controls.[206,210,211]

Although BBB impairment is more commonly associated with VaD than AD, studies in transgenic mice and in humans raise the possibility that BBB dysfunction may be more prevalent in AD than previously believed. Ujiie *et al.*[212] reported that there was increased permeability in the BBB of Tg2576 AD mice in comparison with age-matched controls at 10 months of age as the signs of the disease became manifest. Moreover, the increase in BBB permeability was evident as early as 4 months of age, prior to disease onset and plaque deposition.[212] Further evidence supporting the involvement of the BBB in AD pathogenesis can be seen in studies focusing on the immunization of mice and humans with amyloid peptides and antibodies. In many cases, microhemorrhages occurred in mice after immunization[213,214]; however, the presence of these lesions was antibody-independent. In another Aβ immunization study that focused on a more global degree of BBB status, it was reported that immunization with Aβ appears to repair the damage to the BBB and may even prevent further disease progression.[215] Human studies have also shown an association between BBB impairment and AD. In a small cohort of patients diagnosed with probable AD, BBB impairment was a stable characteristic, as determined by cerebrospinal fluid albumin levels and cerebrospinal fluid/plasma immunoglobulin G levels.[216] Interestingly, this impairment was not associated with vascular factors, ApoE status, or age, and this suggests that BBB impairment in AD may be due to processes distinct from VaD. Recently, Farrall and Wardlaw[217] performed a systematic review of the literature on human clinical studies and found that BBB permeability increases in normal aging and is even more pronounced in patients with dementia and AD. It is important to note that there is significant heterogeneity between studies, and thus more data are needed to clearly resolve this issue.

> *Although blood brain barrier impairment is more commonly associated with vascular dementia than AD, studies in transgenic mice and in humans raise the possibility that blood brain barrier dysfunction may be more prevalent in AD than previously believed.*

OXIDATIVE STRESS

During normal aging and AD, there is a reduction in resting CBF as well as dysfunction in the mechanism that regulates cerebral circulation. This dysfunction results in part from the loss of endothelial mitochondria and a thickening of the vascular basement membrane.[8] Growing evidence appears to implicate oxidative stress as the common factor rendering the brain vulnerable to environmental insults, and it has been shown to play an important role in the pathogenesis of AD (reviewed by Mariani *et al.*[218]). Many of the risk factors that play key roles in AD are associated with vascular oxidative stress; however, whether oxidative stress precedes the onset of AD or exacerbates the pathology is controversial. Oxidative stress, manifested by increased protein oxidation, lipid peroxidation, decreased polyunsaturated fatty acids (PUFAs), and the presence of reactive oxygen species (ROS), is a major characteristic of AD. ROS have long been implicated in the pathogenesis of AD and occur in response to inflammation, injury, and exceedingly low CBF, leading to cell injury and death. There are several sources of vascular ROS, but reduced nicotinamide adenine dinucleotide phosphate (NADPH) oxidase is thought to be one of the main enzymes involved in the production of vascular ROS.[219–221] NADPH is a superoxide-producing enzyme and has been implicated in several oxidative stress conditions including hypertension, and it is significantly activated in AD brains.[219] It has been shown that the presence of excess superoxide ($O_2^{\cdot-}$) radicals in the brains of APP mice is due to the activity of NADPH oxidase, and the inhibition of

> *Growing evidence appears to implicate oxidative stress as the common factor rendering the brain vulnerable to environmental insults, and it has been shown to play an important role in the pathogenesis of AD.*

NADPH activity by either pharmacological inhibitors or the inhibition of the NADPH oxidase complex assembly blocked the production of ROS and cerebrovascular dysfunction induced by Aβ and aging.[90,222] Furthermore, a mouse model lacking the catalytic subunit gp91phox (Nox2) of the enzyme showed a decrease in ROS production, did not show signs of oxidative stress, and appeared to be protected from alterations in the vasculature such as endothelial relaxation and hyperemia.[221,223,224] When APP-overexpressing mice that contained a deletion of Nox2 were compared to APP mice that expressed Nox2, there was no oxidative stress, cerebrovascular dysfunction, or behavioral deficits.[223]

Accumulated oxidative stress affects nitric oxide (NO) function to relax endothelial vasculature, increases vascular endothelial permeability, and further reduces CBF.[8] These are thought to occur because of the reduced bioavailability of NO and the increase in free radicals. When superoxide dismutase (SOD), a potent antioxidant, was incubated in vitro with endothelial cells from aged APP mice, there was a complete restoration of cerebral vascular function.[225] In vivo, when SOD was applied topically to the cerebral cortex of APP transgenic mice, there were no deficits in endothelial function.[226] Moreover, mice expressing both APP and SOD-1 have no endothelial cell deficits.[226] These pronounced oxidation-induced pathological effects in patients with AD and transgenic mice provide evidence that oxidative stress precedes the onset of AD pathology. On the other hand, brains of postmortem AD patients show increased levels of oxidative stress in comparison with non-AD patients, specifically in vascular lesions and mitochondria,[8] and they provide evidence that oxidative stress further attenuates the pathology of AD, whereas levels of antioxidants and antioxidant enzymes are decreased in brains of postmortem AD patients. Levels of PUFAs are also decreased in the brains of AD patients.[8] There is evidence that PUFAs, arachidonic acid, and docosahexaenoic acid are more vulnerable to attack by ROS, and this provides further evidence that oxidative stress exacerbates the pathology of AD. DNA, RNA, and protein oxidation levels are increased in the brains of AD patients, and oxidative stress markers such as AGEs have been found in Aβ plaques and NFTs.[8] Oxidative stress in mitochondria leads to several downstream effects, as mitochondria become less efficient at producing adenosine triphosphate and more efficient at producing ROS; this ultimately results in oxidative stress to the nucleus and cell death.[8]

Amyloid peptides and plaques have been linked to degenerative neurons and to areas high in oxidative stress, and it has been suggested that amyloid is able to induce cerebrovascular dysfunction via oxidative stress mechanisms (see Varadarajan *et al.*[227]). In vitro studies have shown that the application of Aβ peptides to endothelial cells results in the generation of large quantities of $O_2^{\cdot -}$, enhances the rates of apoptosis and necrosis, and prevents the formation of capillary networks.[225,228,229] In mice that overexpress mutated forms of APP, signs of oxidative stress in the vasculature are evident even before plaques have formed.[230] The Aβ-induced oxidative stress impairs cerebrovascular dilatory responses, alters autoregulation and functional hyperemia, and causes cerebral hypometabolism.[224,226,230] Moreover, cerebral vasculature dysfunction can be rescued in aged APP mice treated with antioxidants.[231] Another model names Aβ peptide bound to redox metal ions as the culprit of neurotoxicity found in AD. Varadarajan *et al.*[227] proposed that small soluble Aβ aggregates insert into the membranes of neurons or glia and generate oxygenated free radicals that cause protein oxidation and lipid peroxidation. This, in turn, causes membrane disruption, which leads to cellular dysfunction, including perturbation of calcium homeostasis, transporter function, and activation of apoptotic signaling pathways.

Oxidative stress is associated with negative pathology related to a reduction in CBF, detriments to NO and endothelial vasculature, Aβ plaques, and ultimately cell death. Because of the vast role of oxidative stress in preceding or exacerbating AD pathology, with the help of antioxidants, it can act as a major therapeutic target in the onset and pathology of AD.

CONCLUSION

At first glance, the relationship between vascular risk factors and AD may appear contradictory, as vascular risk factors and the presence of vascular disease were considered exclusion criteria for the clinical diagnosis of AD. However, recent studies have suggested that these co-occurrences, both common in the elderly and believed to occur by chance, have more pathological significance. Studies have suggested that microvascular disorder and dysfunction can contribute to cognitive decline and pathology associated with AD in a synergistic manner and can exacerbate the clinical symptomatology of AD. Although there are still conflicting views on the evidence put forth from the many epidemiological studies, it is agreed that there does indeed appear to be a relationship between many vascular risk factors, vascular dysfunction, and cognitive decline.

These data have led to the vascular hypothesis for AD, which proposes that dysfunction of the neurovasculature and nonneuronal neighboring cells contributes to the pathogenesis of dementia and AD. However, as the exact relationships between (and potentially shared mechanisms of) vascular risk factors, vascular dysfunction, and neuronal degeneration remain poorly understood, it is difficult to state definitive conclusions. More prospective studies on human patients with longer follow-up periods as well as studies in transgenic mice are needed to resolve this issue. Further exploration of the roles of vascular risk factors and vascular dysfunction in the pathogenesis of AD and cognitive decline may provide a better understanding of the molecular mechanisms and sharper therapeutic targets for intervention in the future.

ACKNOWLEDGMENT

The authors thank members of the Hof and Gandy laboratories for their help and discussion. This work was supported by grants AG05138, AG02219, and AG10491 from the National Institutes of Health.

DISCLOSURES

Potential conflict of interest: Samuel Gandy is currently a consultant to Diagenic, Pfizer/Janssen, and Amicus and was a consultant to Epix in the past. He also has an Investigator-initiated grant from Forest.

REFERENCES

1. Takeda S, Sato N, Ogihara T, Morishita R. The renin-angiotensin system, hypertension and cognitive dysfunction in Alzheimer's disease: new therapeutic potential. *Front Biosci* 2008; 13: 2253–2265.
2. Ferri CP, Prince M, Brayne C, et al. Global prevalence of dementia: a Delphi consensus study. *Lancet* 2005; 366: 2112–2117.
3. Whitehouse PJ, Sciulli CG, Mason RM. Dementia drug development: use of information systems to harmonize global drug development. *Psychopharmacol Bull* 1997; 33: 129–133.
4. Hardy J. A hundred years of Alzheimer's disease research. *Neuron* 2006; 52: 3–13.
5. Kivipelto M, Helkala EL, Hanninen T, et al. Midlife vascular risk factors and late-life mild cognitive impairment: A population-based study. *Neurology* 2001; 56: 1683–1689.
6. Ott A, Stolk RP, van Harskamp F, et al. Diabetes mellitus and the risk of dementia: The Rotterdam Study. *Neurology* 1999; 53: 1937–1942.
7. Ott A, Slooter AJ, Hofman A, et al. Smoking and risk of dementia and Alzheimer's disease in a population-based cohort study: the Rotterdam Study. *Lancet* 1998; 351: 1840–1843.
8. Zhu X, Smith MA, Honda K, et al. Vascular oxidative stress in Alzheimer disease. *J Neurol Sci* 2007; 257: 240–246.
9. Dede DS, Yavuz B, Yavuz BB, et al. Assessment of endothelial function in Alzheimer's disease: is Alzheimer's disease a vascular disease?. *J Am Geriatr Soc* 2007; 55: 1613–1617.
10. Jellinger KA, Attems J. Incidence of cerebrovascular lesions in Alzheimer's disease: a postmortem study. *Acta Neuropathol* 2003; 105: 14–17.
11. Jellinger KA. The enigma of vascular cognitive disorder and vascular dementia. *Acta Neuropathol* 2007; 113: 349–388.
12. Tomlinson BE, Blessed G, Roth M. Observations on the brains of demented old people. *J Neurol Sci* 1970; 11: 205–242.
13. del Ser T, Bermejo F, Portera A, et al. Vascular dementia. A clinicopathological study. *J Neurol Sci* 1990; 96: 1–17.
14. Esiri MM, Wilcock GK, Morris JH. Neuropathological assessment of the lesions of significance in vascular dementia. *J Neurol Neurosurg Psychiatry* 1997; 63: 749–753.
15. Snowdon DA, Greiner LH, Mortimer JA, et al. Brain infarction and the clinical expression of Alzheimer disease. *JAMA* 1997; 277: 813–817.
16. Esiri MM, Nagy Z, Smith MZ, et al. Cerebrovascular disease and threshold for dementia in the early stages of Alzheimer's disease. *Lancet* 1999; 354: 919–920.
17. Vinters HV, Ellis WG, Zarow C, et al. Neuropathologic substrates of ischemic vascular dementia. *J Neuropathol Exp Neurol* 2000; 59: 931–945.
18. Kövari E, Gold G, Herrmann FR, et al. Lewy body densities in the entorhinal and anterior cingulate cortex predict cognitive deficits in Parkinson's disease. *Acta Neuropathol* 2003; 106: 83–88.
19. Gold G, Kövari E, Corte G, et al. Clinical validity of A beta-protein deposition staging in brain aging and Alzheimer disease. *J Neuropathol Exp Neurol* 2001; 60: 946–952.
20. Giannakopoulos P, Herrmann FR, Bussière T, et al. Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology* 2003; 60: 1495–1500.
21. Vermeer SE, Prins ND, den Heijer T, et al. Silent brain infarcts and the risk of dementia and cognitive decline. *N Engl J Med* 2003; 348: 1215–1222.
22. Roman GC, Erkinjuntti T, Wallin A, et al. Subcortical ischaemic vascular dementia. *Lancet Neurol* 2002; 1: 426–436.
23. Gold G, Kövari E, Herrmann FR, et al. Cognitive consequences of thalamic, basal ganglia, and deep white matter lacunes in brain aging and dementia. *Stroke* 2005; 36: 1184–1188.
24. Gold G, Kövari E, Hof PR, et al. Sorting out the clinical consequences of ischemic lesions in brain aging: a clinicopathological approach. *J Neurol Sci* 2007; 257: 17–22.
25. Roman GC, Royall DR. Executive control function: a rational basis for the diagnosis of vascular dementia. *Alzheimer Dis Assoc Disord* 1999; 13(suppl 3): S69–S80.

26. Jellinger KA, Mitter-Ferstl E. The impact of cerebrovascular lesions in Alzheimer disease–a comparative autopsy study. *J Neurol* 2003; 250: 1050–1055.
27. Gold G, Giannakopoulos P, Herrmann FR, et al. Identification of Alzheimer and vascular lesion thresholds for mixed dementia. *Brain* 2007; 130: 2830–2836.
28. Bailey TL, Rivara CB, Rocher AB, Hof PR. The nature and effects of cortical microvascular pathology in aging and Alzheimer's disease. *Neurol Res* 2004; 26: 573–578.
29. Buée L, Hof PR, Bouras C, et al. Pathological alterations of the cerebral microvasculature in Alzheimer's disease and related dementing disorders. *Acta Neuropathol* 1994; 87: 469–480.
30. Attems J. Sporadic cerebral amyloid angiopathy: pathology, clinical implications, and possible pathomechanisms. *Acta Neuropathol* 2005; 110: 345–359.
31. Herzig MC, Van Nostrand WE, Jucker M. Mechanism of cerebral beta-amyloid angiopathy: murine and cellular models. *Brain Pathol* 2006; 16: 40–54.
32. Ferrer I, Boada Rovira M, Sanchez Guerra ML, et al. Neuropathology and pathogenesis of encephalitis following amyloid-beta immunization in Alzheimer's disease. *Brain Pathol* 2004; 14: 11–20.
33. Kawai M, Kalaria RN, Harik SI, Perry G. The relationship of amyloid plaques to cerebral capillaries in Alzheimer's disease. *Am J Pathol* 1990; 137: 1435–1446.
34. Revesz T, Holton JL, Lashley T, et al. Sporadic and familial cerebral amyloid angiopathies. *Brain Pathol* 2002; 12: 343–357.
35. Johnson-Wood K, Lee M, Motter R, et al. Amyloid precursor protein processing and A beta42 deposition in a transgenic mouse model of Alzheimer disease. *Proc Natl Acad Sci U S A* 1997; 94: 1550–1555.
36. Calhoun ME, Burgermeister P, Phinney AL, et al. Neuronal overexpression of mutant amyloid precursor protein results in prominent deposition of cerebrovascular amyloid. *Proc Natl Acad Sci U S A* 1999; 96: 14 088–14 093.
37. Van Dorpe J, Smeijers L, Dewachter I, et al. Prominent cerebral amyloid angiopathy in transgenic mice overexpressing the london mutant of human APP in neurons. *Am J Pathol* 2000; 157: 1283–1298.
38. Winkler DT, Bondolfi L, Herzig MC, et al. Spontaneous hemorrhagic stroke in a mouse model of cerebral amyloid angiopathy. *J Neurosci* 2001; 21: 1619–1627.
39. Fryer JD, Taylor JW, DeMattos RB, et al. Apolipoprotein E markedly facilitates age-dependent cerebral amyloid angiopathy and spontaneous hemorrhage in amyloid precursor protein transgenic mice. *J Neurosci* 2003; 23: 7889–7896.
40. Davis J, Xu F, Deane R, et al. Early-onset and robust cerebral microvascular accumulation of amyloid beta-protein in transgenic mice expressing low levels of a vasculotropic Dutch/Iowa mutant form of amyloid beta-protein precursor. *J Biol Chem* 2004; 279: 20 296–20 306.
41. Herzig MC, Winkler DT, Burgermeister P, et al. Abeta is targeted to the vasculature in a mouse model of hereditary cerebral hemorrhage with amyloidosis. *Nat Neurosci* 2004; 7: 954–960.
42. Kumar-Singh S, Pirici D, McGowan E, et al. Dense-core plaques in Tg2576 and PSAPP mouse models of Alzheimer's disease are centered on vessel walls. *Am J Pathol* 2005; 167: 527–543.
43. Miao J, Xu F, Davis J, et al. Cerebral microvascular amyloid beta protein deposition induces vascular degeneration and neuroinflammation in transgenic mice expressing human vasculotropic mutant amyloid beta precursor protein. *Am J Pathol* 2005; 167: 505–515.
44. Fryer JD, Simmons K, Parsadanian M, et al. Human apolipoprotein E4 alters the amyloid-beta 40:42 ratio and promotes the formation of cerebral amyloid angiopathy in an amyloid precursor protein transgenic model. *J Neurosci* 2005; 25: 2803–2810.
45. Mahley RW. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science* 1988; 240: 622–630.
46. Mahley RW, Huang Y. Apolipoprotein (apo) E4 and Alzheimer's disease: unique conformational and biophysical properties of apoE4 can modulate neuropathology. *Acta Neurol Scand Suppl* 2006; 185: 8–14.
47. Panza F, D'Introno A, Colacicco AM, et al. Lipid metabolism in cognitive decline and dementia. *Brain Res Rev* 2006; 51: 275–292.
48. Donahue JE, Johanson CE. Apolipoprotein E, amyloid-beta, and blood-brain barrier permeability in Alzheimer disease. *J Neuropathol Exp Neurol* 2008; 67: 261–270.
49. Saunders AM, Schmader K, Breitner JC, et al. Apolipoprotein E epsilon 4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. *Lancet* 1993; 342: 710–711.
50. Blacker D, Haines JL, Rodes L, et al. ApoE-4 and age at onset of Alzheimer's disease: the NIMH genetics initiative. *Neurology* 1997; 48: 139–147.
51. Davignon J, Gregg RE, Sing CF. Apolipoprotein E polymorphism and atherosclerosis. *Arteriosclerosis* 1988; 8: 1–21.
52. Kobori S, Nakamura N, Uzawa H, Shichiri M. Influence of apolipoprotein E polymorphism on plasma lipid and apolipoprotein levels, and clinical characteristics of type III hyperlipoproteinemia due to apolipoprotein E phenotype E2/2 in Japan. *Atherosclerosis* 1988; 69: 81–88.
53. Kessler C, Spitzer C, Stauske D, et al. The apolipoprotein E and beta-fibrinogen G/A-455 gene polymorphisms are associated with ischemic stroke involving large-vessel disease. *Arterioscler Thromb Vasc Biol* 1997; 17: 2880–2884.
54. McCarron MO, Delong D, Alberts MJ. APOE genotype as a risk factor for ischemic cerebrovascular disease: a meta-analysis. *Neurology* 1999; 53: 1308–1311.
55. Tanskanen M, Lindsberg PJ, Tienari PJ, et al. Cerebral amyloid angiopathy in a 95+ cohort: complement activation and apolipoprotein E (ApoE) genotype. *Neuropathol Appl Neurobiol* 2005; 31: 589–599.
56. Peila R, Rodriguez BL, Launer LJ. Type 2 diabetes, APOE gene, and the risk for dementia and related pathologies: The Honolulu-Asia Aging Study. *Diabetes* 2002; 51: 1256–1262.
57. Slooter AJ, Tang MX, Van Duijn CM, et al. Apolipoprotein E epsilon4 and the risk of dementia with stroke. A population-based investigation. *JAMA* 1997; 277: 818–821.

58. Cosentino S, Scarmeas N, Helzner E, et al. APOE epsilon 4 allele predicts faster cognitive decline in mild Alzheimer disease. *Neurology* 2008; 70: 1842–1849.
59. Xu F, Vitek MP, Colton CA, et al. Human apolipoprotein E redistributes fibrillar amyloid deposition in Tg-SwDI mice. *J Neurosci* 2008; 28: 5312–5320.
60. Holtzman DM, Bales KR, Tenkova T, et al. Apolipoprotein E isoform-dependent amyloid deposition and neuritic degeneration in a mouse model of Alzheimer's disease. *Proc Natl Acad Sci U S A* 2000; 97: 2892–2897.
61. Holtzman DM, Fagan AM, Mackey B, et al. Apolipoprotein E facilitates neuritic and cerebrovascular plaque formation in an Alzheimer's disease model. *Ann Neurol* 2000; 47: 739–747.
62. DeMattos RB, Cirrito JR, Parsadanian M, et al. ApoE and clusterin cooperatively suppress Abeta levels and deposition: evidence that ApoE regulates extracellular Abeta metabolism in vivo. *Neuron* 2004; 41: 193–202.
63. Jiang Q, Lee CY, Mandrekar S, et al. ApoE promotes the proteolytic degradation of Abeta. *Neuron* 2008; 58: 681–693.
64. Deane R, Sagare A, Hamm K, et al. apoE isoform-specific disruption of amyloid beta peptide clearance from mouse brain. *J Clin Invest* 2008; 118: 4002–4013.
65. Poirier J. Apolipoprotein E and cholesterol metabolism in the pathogenesis and treatment of Alzheimer's disease. *Trends Mol Med* 2003; 9: 94–101.
66. Dannenberg AL, Garrison RJ, Kannel WB. Incidence of hypertension in the Framingham Study. *Am J Public Health* 1988; 78: 676–679.
67. Messerli FH, Williams B, Ritz E. Essential hypertension. *Lancet* 2007; 370: 591–603.
68. Skoog I, Gustafson D. Update on hypertension and Alzheimer's disease. *Neurol Res* 2006; 28: 605–611.
69. Paglieri C, Bisbocci D, Caserta M, et al. Hypertension and cognitive function. *Clin Exp Hypertens* 2008; 30: 701–710.
70. Elias MF, D'Agostino RB, Elias PK, Wolf PA. Neuropsychological test performance, cognitive functioning, blood pressure, and age: the Framingham Heart Study. *Exp Aging Res* 1995; 21: 369–391.
71. Yoshitake T, Kiyohara Y, Kato I, et al. Incidence and risk factors of vascular dementia and Alzheimer's disease in a defined elderly Japanese population: the Hisayama Study. *Neurology* 1995; 45: 1161–1168.
72. Skoog I, Lernfelt B, Landahl S, et al. 15-year longitudinal study of blood pressure and dementia. *Lancet* 1996; 347: 1141–1145.
73. Launer LJ, Ross GW, Petrovitch H, et al. Midlife blood pressure and dementia: the Honolulu-Asia aging study. *Neurobiol Aging* 2000; 21: 49–55.
74. Posner HB, Tang MX, Luchsinger J, et al. The relationship of hypertension in the elderly to AD, vascular dementia, and cognitive function. *Neurology* 2002; 58: 1175–1181.
75. Farmer ME, Kittner SJ, Abbott RD, et al. Longitudinally measured blood pressure, antihypertensive medication use, and cognitive performance: the Framingham Study. *J Clin Epidemiol* 1990; 43: 475–480.
76. Elias MF, Wolf PA, D'Agostino RB, et al. Untreated blood pressure level is inversely related to cognitive functioning: the Framingham Study. *Am J Epidemiol* 1993; 138: 353–364.
77. Sparks DL, Scheff SW, Liu H, et al. Increased incidence of neurofibrillary tangles (NFT) in non-demented individuals with hypertension. *J Neurol Sci* 1995; 131: 162–169.
78. Petrovitch H, White LR, Izmirilian G, et al. Midlife blood pressure and neuritic plaques, neurofibrillary tangles, and brain weight at death: the HAAS. Honolulu-Asia aging Study. *Neurobiol Aging* 2000; 21: 57–62.
79. Guo Z, Viitanen M, Fratiglioni L, Winblad B. Low blood pressure and dementia in elderly people: the Kungsholmen project. *BMJ* 1996; 312: 805–808.
80. Zhu L, Fratiglioni L, Guo Z, et al. Association of stroke with dementia, cognitive impairment, and functional disability in the very old: a population-based study. *Stroke* 1998; 29: 2094–2099.
81. Morris MC, Scherr PA, Hebert LE, et al. Association of incident Alzheimer disease and blood pressure measured from 13 years before to 2 years after diagnosis in a large community study. *Arch Neurol* 2001; 58: 1640–1646.
82. Mattila K, Haavisto M, Rajala S, Heikinheimo R. Blood pressure and five year survival in the very old. *Br Med J (Clin Res Ed)* 1988; 296: 887–889.
83. De Buyzere M, Clement DL, Duprez D. Chronic low blood pressure: a review. *Cardiovasc Drugs Ther* 1998; 12: 29–35.
84. Moretti R, Torre P, Antonello RM, et al. Risk factors for vascular dementia: hypotension as a key point. *Vasc Health Risk Manag* 2008; 4: 395–402.
85. Maule S, Caserta M, Bertello C, et al. Cognitive decline and low blood pressure: the other side of the coin. *Clin Exp Hypertens* 2008; 30: 711–719.
86. Qiu C, von Strauss E, Fastbom J, et al. Low blood pressure and risk of dementia in the Kungsholmen project: a 6-year follow-up study. *Arch Neurol* 2003; 60: 223–228.
87. Guo Z, Viitanen M, Winblad B, Fratiglioni L. Low blood pressure and incidence of dementia in a very old sample: dependent on initial cognition. *J Am Geriatr Soc* 1999; 47: 723–726.
88. Morris MC, Scherr PA, Hebert LE, et al. The cross-sectional association between blood pressure and Alzheimer's disease in a biracial community population of older persons. *J Gerontol A Biol Sci Med Sci* 2000; 55: M130–M136.
89. Ruitenberg A, Skoog I, Ott A, et al. Blood pressure and risk of dementia: results from the Rotterdam study and the Gothenburg H-70 Study. *Dement Geriatr Cogn Disord* 2001; 12: 33–39.
90. Iadecola C, Davisson RL. Hypertension and cerebrovascular dysfunction. *Cell Metab* 2008; 7: 476–484.
91. Kennelly SP, Lawlor BA, Kenny RA. Blood pressure and the risk for dementia: a double edged sword. *Ageing Res Rev* 2009; 8: 61–70.
92. Burke WJ, Coronado PG, Schmitt CA, et al. Blood pressure regulation in Alzheimer's disease. *J Auton Nerv Syst* 1994; 48: 65–71.
93. Qiu C, von Strauss E, Winblad B, Fratiglioni L. Decline in blood pressure over time and risk of dementia: a longitudinal study from the Kungsholmen project. *Stroke* 2004; 35: 1810–1815.

94. Wolozin B. Cholesterol, statins and dementia. *Curr Opin Lipidol* 2004; 15: 667–672.
95. Duron E, Hanon O. Vascular risk factors, cognitive decline, and dementia. *Vasc Health Risk Manag* 2008; 4: 363–381.
96. Reid PC, Urano Y, Kodama T, Hamakubo T. Alzheimer's disease: cholesterol, membrane rafts, isoprenoids and statins. *J Cell Mol Med* 2007; 11: 383–392.
97. Simons M, Keller P, De Strooper B, et al. Cholesterol depletion inhibits the generation of beta-amyloid in hippocampal neurons. *Proc Natl Acad Sci U S A* 1998; 95: 6460–6464.
98. Wahrle S, Das P, Nyborg AC, et al. Cholesterol-dependent gamma-secretase activity in buoyant cholesterol-rich membrane microdomains. *Neurobiol Dis* 2002; 9: 11–23.
99. Ehehalt R, Keller P, Haass C, et al. Amyloidogenic processing of the Alzheimer beta-amyloid precursor protein depends on lipid rafts. *J Cell Biol* 2003; 160: 113–123.
100. Bodovitz S, Klein WL. Cholesterol modulates alpha-secretase cleavage of amyloid precursor protein. *J Biol Chem* 1996; 271: 4436–4440.
101. Kojro E, Gimpl G, Lammich S, et al. Low cholesterol stimulates the nonamyloidogenic pathway by its effect on the alpha -secretase ADAM 10. *Proc Natl Acad Sci U S A* 2001; 98: 5815–5820.
102. Fassbender K, Simons M, Bergmann C, et al. Simvastatin strongly reduces levels of Alzheimer's disease beta -amyloid peptides Abeta 42 and Abeta 40 in vitro and in vivo. *Proc Natl Acad Sci U S A* 2001; 98: 5856–5861.
103. Guardia-Laguarta C, Coma M, Pera M, et al. Mild cholesterol depletion reduces amyloid-beta production by impairing APP trafficking to the cell surface. *J Neurochem* 2009; 110: 220–230.
104. Wada S, Morishima-Kawashima M, Qi Y, et al. Gamma-secretase activity is present in rafts but is not cholesterol-dependent. *Biochemistry* 2003; 42: 13 977–13 986.
105. Ostrowski SM, Wilkinson BL, Golde TE, Landreth G. Statins reduce amyloid-beta production through inhibition of protein isoprenylation. *J Biol Chem* 2007; 282: 26 832–26 844.
106. Xiong H, Callaghan D, Jones A, et al. Cholesterol retention in Alzheimer's brain is responsible for high beta- and gamma-secretase activities and Abeta production. *Neurobiol Dis* 2008; 29: 422–437.
107. Racchi M, Baetta R, Salvietti N, et al. Secretory processing of amyloid precursor protein is inhibited by increase in cellular cholesterol content. *Biochem J* 1997; 322(pt 3): 893–898.
108. Sparks DL, Scheff SW, Hunsaker JC III, et al. Induction of Alzheimer-like beta-amyloid immunoreactivity in the brains of rabbits with dietary cholesterol. *Exp Neurol* 1994; 126: 88–94.
109. Sparks DL, Frank PG, Braschi S, et al. Effect of apolipoprotein A-I lipidation on the formation and function of pre-beta and alpha-migrating LpA-I particles. *Biochemistry* 1999; 38: 1727–1735.
110. Tamboli IY, Prager K, Thal DR, et al. Loss of gamma-secretase function impairs endocytosis of lipoprotein particles and membrane cholesterol homeostasis. *J Neurosci* 2008; 28: 12 097–12 106.
111. Refolo LM, Malester B, LaFrancois J, et al. Hypercholesterolemia accelerates the Alzheimer's amyloid pathology in a transgenic mouse model. *Neurobiol Dis* 2000; 7: 321–331.
112. Notkola IL, Sulkava R, Pekkanen J, et al. Serum total cholesterol, apolipoprotein E epsilon 4 allele, and Alzheimer's disease. *Neuroepidemiology* 1998; 17: 14–20.
113. Kivipelto M, Helkala EL, Laakso MP, et al. Apolipoprotein E epsilon4 allele, elevated midlife total cholesterol level, and high midlife systolic blood pressure are independent risk factors for late-life Alzheimer disease. *Ann Intern Med* 2002; 137: 149–155.
114. Tan ZS, Seshadri S, Beiser A, et al. Plasma total cholesterol level as a risk factor for Alzheimer disease: the Framingham Study. *Arch Intern Med* 2003; 163: 1053–1057.
115. Beckett N, Nunes M, Bulpitt C. Is it advantageous to lower cholesterol in the elderly hypertensive?. *Cardiovasc Drugs Ther* 2000; 14: 397–405.
116. Kuusisto J, Koivisto K, Mykkanen L, et al. Association between features of the insulin resistance syndrome and Alzheimer's disease independently of apolipoprotein E4 phenotype: cross sectional population based study. *BMJ* 1997; 315: 1045–1049.
117. Romas SN, Tang MX, Berglund L, Mayeux R. APOE genotype, plasma lipids, lipoproteins, and AD in community elderly. *Neurology* 1999; 53: 517–521.
118. Evans RM, Emsley CL, Gao S, et al. Serum cholesterol, APOE genotype, and the risk of Alzheimer's disease: a population-based study of African Americans. *Neurology* 2000; 54: 240–242.
119. Jick H, Zornberg GL, Jick SS, et al. Statins and the risk of dementia. *Lancet* 2000; 356: 1627–1631.
120. Wolozin B, Kellman W, Ruosseau P, et al. Decreased prevalence of Alzheimer disease associated with 3-hydroxy-3-methyglutaryl coenzyme A reductase inhibitors. *Arch Neurol* 2000; 57: 1439–1443.
121. Rockwood K, Kirkland S, Hogan DB, et al. Use of lipid-lowering agents, indication bias, and the risk of dementia in community-dwelling elderly people. *Arch Neurol* 2002; 59: 223–227.
122. Johnson-Anuna LN, Eckert GP, Keller JH, et al. Chronic administration of statins alters multiple gene expression patterns in mouse cerebral cortex. *J Pharmacol Exp Ther* 2005; 312: 786–793.
123. Shepherd J, Blauw GJ, Murphy MB, et al. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet* 2002; 360: 1623–1630.
124. Biessels GJ, Staekenborg S, Brunner E, et al. Risk of dementia in diabetes mellitus: a systematic review. *Lancet Neurol* 2006; 5: 64–74.
125. Brands AM, Biessels GJ, de Haan EH, et al. The effects of type 1 diabetes on cognitive performance: a meta-analysis. *Diabetes Care* 2005; 28: 726–735.
126. Allen KV, Frier BM, Strachan MW. The relationship between type 2 diabetes and cognitive dysfunction: longitudinal studies and their methodological limitations. *Eur J Pharmacol* 2004; 490: 169–175.
127. Bucht G, Adolfsson R, Lithner F, Winblad B. Changes in blood glucose and insulin secretion in patients with senile dementia of Alzheimer type. *Acta Med Scand* 1983; 213: 387–392.

128. Wolf-Klein GP, Siverstone FA, Brod MS, et al. Are Alzheimer patients healthier?. *J Am Geriatr Soc* 1988; 36: 219–224.
129. Nielson KA, Nolan JH, Berchtold NC, et al. Apolipoprotein-E genotyping of diabetic dementia patients: is diabetes rare in Alzheimer's disease?. *J Am Geriatr Soc* 1996; 44: 897–904.
130. Leibson CL, Rocca WA, Hanson VA, et al. Risk of dementia among persons with diabetes mellitus: a population-based cohort study. *Am J Epidemiol* 1997; 145: 301–308.
131. MacKnight C, Rockwood K, Awalt E, McDowell I. Diabetes mellitus and the risk of dementia, Alzheimer's disease and vascular cognitive impairment in the Canadian Study of Health and Aging. *Dement Geriatr Cogn Disord* 2002; 14: 77–83.
132. Janson J, Laedtke T, Parisi JE, et al. Increased risk of type 2 diabetes in Alzheimer disease. *Diabetes* 2004; 53: 474–481.
133. Pasquier F, Boulogne A, Leys D, Fontaine P. Diabetes mellitus and dementia. *Diabetes Metab* 2006; 32: 403–414.
134. Kalmijn S, Foley D, White L, et al. Metabolic cardiovascular syndrome and risk of dementia in Japanese-American elderly men. The Honolulu-Asia aging study. *Arterioscler Thromb Vasc Biol* 2000; 20: 2255–2260.
135. Yaffe K, Kanaya A, Lindquist K, et al. The metabolic syndrome, inflammation, and risk of cognitive decline. *JAMA* 2004; 292: 2237–2242.
136. Whitmer RA, Gunderson EP, Barrett-Connor E, et al. Obesity in middle age and future risk of dementia: a 27 year longitudinal population based study. *BMJ* 2005; 330: 1360.
137. Gispen WH, Biessels GJ. Cognition and synaptic plasticity in diabetes mellitus. *Trends Neurosci* 2000; 23: 542–549.
138. Kumari MV, Hiramatsu M, Ebadi M. Free radical scavenging actions of hippocampal metallothionein isoforms and of antimetallothioneins: an electron spin resonance spectroscopic study. *Cell Mol Biol (Noisy-le-grand)* 2000; 46: 627–636.
139. Brownlee M. Biochemistry and molecular cell biology of diabetic complications. *Nature* 2001; 414: 813–820.
140. Degenhardt TP, Thorpe SR, Baynes JW. Chemical modification of proteins by methylglyoxal. *Cell Mol Biol (Noisy-le-grand)* 1998; 44: 1139–1145.
141. Thornalley PJ. The glyoxalase system: new developments towards functional characterization of a metabolic pathway fundamental to biological life. *Biochem J* 1990; 269: 1–11.
142. Wells-Knecht KJ, Zyzak DV, Litchfield JE, et al. Mechanism of autoxidative glycosylation: identification of glyoxal and arabinose as intermediates in the autoxidative modification of proteins by glucose. *Biochemistry* 1995; 34: 3702–3709.
143. Giardino I, Edelstein D, Brownlee M. Nonenzymatic glycosylation in vitro and in bovine endothelial cells alters basic fibroblast growth factor activity. A model for intracellular glycosylation in diabetes. *J Clin Invest* 1994; 94: 110–117.
144. Shinohara M, Thornalley PJ, Giardino I, et al. Overexpression of glyoxalase-I in bovine endothelial cells inhibits intracellular advanced glycation endproduct formation and prevents hyperglycemia-induced increases in macromolecular endocytosis. *J Clin Invest* 1998; 101: 1142–1147.
145. Huijberts MS, Wolffenbuttel BH, Boudier HA, et al. Aminoguanidine treatment increases elasticity and decreases fluid filtration of large arteries from diabetic rats. *J Clin Invest* 1993; 92: 1407–1411.
146. Yan SD, Schmidt AM, Anderson GM, et al. Enhanced cellular oxidant stress by the interaction of advanced glycation end products with their receptors/binding proteins. *J Biol Chem* 1994; 269: 9889–9897.
147. Schmidt AM, Hori O, Chen JX, et al. Advanced glycation endproducts interacting with their endothelial receptor induce expression of vascular cell adhesion molecule-1 (VCAM-1) in cultured human endothelial cells and in mice. A potential mechanism for the accelerated vasculopathy of diabetes. *J Clin Invest* 1995; 96: 1395–1403.
148. Abordo EA, Westwood ME, Thornalley PJ. Synthesis and secretion of macrophage colony stimulating factor by mature human monocytes and human monocytic THP-1 cells induced by human serum albumin derivatives modified with methylglyoxal and glucose-derived advanced glycation endproducts. *Immunol Lett* 1996; 53: 7–13.
149. Lu M, Kuroki M, Amano S, et al. Advanced glycation end products increase retinal vascular endothelial growth factor expression. *J Clin Invest* 1998; 101: 1219–1224.
150. Peppa M, Vlassara H. Advanced glycation end products and diabetic complications: a general overview. *Hormones (Athens)* 2005; 4: 28–37.
151. Luchsinger JA, Reitz C, Honig LS, et al. Aggregation of vascular risk factors and risk of incident Alzheimer disease. *Neurology* 2005; 65: 545–551.
152. Baron AD. Hemodynamic actions of insulin. *Am J Physiol* 1994; 267: E187–E202.
153. Ling X, Martins RN, Racchi M, et al. Amyloid beta antagonizes insulin promoted secretion of the amyloid beta protein precursor. *J Alzheimers Dis* 2002; 4: 369–374.
154. Farris W, Mansourian S, Chang Y, et al. Insulin-degrading enzyme regulates the levels of insulin, amyloid beta-protein, and the beta-amyloid precursor protein intracellular domain in vivo. *Proc Natl Acad Sci U S A* 2003; 100: 4162–4167.
155. Gerozissis K. Brain insulin and feeding: a bi-directional communication. *Eur J Pharmacol* 2004; 490: 59–70.
156. Banks WA. The source of cerebral insulin. *Eur J Pharmacol* 2004; 490: 5–12.
157. Steen E, Terry BM, Rivera EJ, et al. Impaired insulin and insulin-like growth factor expression and signaling mechanisms in Alzheimer's disease–is this type 3 diabetes?. *J Alzheimers Dis* 2005; 7: 63–80.
158. Bondy CA, Cheng CM. Signaling by insulin-like growth factor 1 in brain. *Eur J Pharmacol* 2004; 490: 25–31.
159. Zhao WQ, Alkon DL. Role of insulin and insulin receptor in learning and memory. *Mol Cell Endocrinol* 2001; 177: 125–134.
160. Frolich L, Blum-Degen D, Bernstein HG, et al. Brain insulin and insulin receptors in aging and sporadic Alzheimer's disease. *J Neural Transm* 1998; 105: 423–438.

161. Caccamo A, Oddo S, Sugarman MC, et al. Age- and region-dependent alterations in Abeta-degrading enzymes: implications for Abeta-induced disorders. *Neurobiol Aging* 2005; 26: 645–654.
162. Carro E, Trejo JL, Gomez-Isla T, et al. Serum insulin-like growth factor I regulates brain amyloid-beta levels. *Nat Med* 2002; 8: 1390–1397.
163. Selkoe DJ. Clearing the brain's amyloid cobwebs. *Neuron* 2001; 32: 177–180.
164. Gasparini L, Xu H. Potential roles of insulin and IGF-I in Alzheimer's disease. *Trends Neurosci* 2003; 26: 404–406.
165. Ho L, Qin W, Pompl PN, et al. Diet-induced insulin resistance promotes amyloidosis in a transgenic mouse model of Alzheimer's disease. *FASEB J* 2004; 18: 902–904.
166. Graves AB, Van Duijn CM, Chandra V, et al. Alcohol and tobacco consumption as risk factors for Alzheimer's disease: a collaborative re-analysis of case-control studies. EURODEM Risk Factors Research Group. *Int J Epidemiol* 1991; 20(suppl 2): S48–S57.
167. Van Duijn CM, Hofman A. Relation between nicotine intake and Alzheimer's disease. *BMJ* 1991; 302: 1491–1494.
168. Brenner DE, Kukull WA, Van Belle G, et al. Relationship between cigarette smoking and Alzheimer's disease in a population-based case-control study. *Neurology* 1993; 43: 293–300.
169. Shinton R, Beevers G. Meta-analysis of relation between cigarette smoking and stroke. *BMJ* 1989; 298: 789–794.
170. Cruickshank JM, Neil-Dwyer G, Dorrance DE, et al. Acute effects of smoking on blood pressure and cerebral blood flow. *J Hum Hypertens* 1989; 3: 443–449.
171. Hebert LE, Scherr PA, Beckett LA, et al. Relation of smoking and alcohol consumption to incident Alzheimer's disease. *Am J Epidemiol* 1992; 135: 347–355.
172. Broe GA, Creasey H, Jorm AF, et al. Health habits and risk of cognitive impairment and dementia in old age: a prospective study on the effects of exercise, smoking and alcohol consumption. *Aust N Z J Public Health* 1998; 22: 621–623.
173. Doll R, Peto R, Boreham J, Sutherland I. Smoking and dementia in male British doctors: prospective study. *BMJ* 2000; 320: 1097–1102.
174. Galanis DJ, Petrovitch H, Launer LJ, et al. Smoking history in middle age and subsequent cognitive performance in elderly Japanese-American men. The Honolulu-Asia Aging Study. *Am J Epidemiol* 1997; 145: 507–515.
175. Merchant C, Tang MX, Albert S, et al. The influence of smoking on the risk of Alzheimer's disease. *Neurology* 1999; 52: 1408–1412.
176. Ott A, Andersen K, Dewey ME, et al. Effect of smoking on global cognitive function in nondemented elderly. *Neurology* 2004; 62: 920–924.
177. Fratiglioni L, Wang HX. Smoking and Parkinson's and Alzheimer's disease: review of the epidemiological studies. *Behav Brain Res* 2000; 113: 117–120.
178. Launer LJ, Andersen K, Dewey ME, et al. Rates and risk factors for dementia and Alzheimer's disease: results from EURODEM pooled analyses. EURODEM Incidence Research Group and Work Groups. European Studies of Dementia. *Neurology* 1999; 52: 78–84.
179. Wang HX, Fratiglioni L, Frisoni GB, et al. Smoking and the occurrence of Alzheimer's disease: cross-sectional and longitudinal data in a population-based study. *Am J Epidemiol* 1999; 149: 640–644.
180. Aggarwal NT, Bienias JL, Bennett DA, et al. The relation of cigarette smoking to incident Alzheimer's disease in a biracial urban community population. *Neuroepidemiology* 2006; 26: 140–146.
181. Reitz C, den Heijer T, Van Duijn C, et al. Relation between smoking and risk of dementia and Alzheimer disease: the Rotterdam Study. *Neurology* 2007; 69: 998–1005.
182. Kihara T, Shimohama S, Sawada H, et al. Nicotinic receptor stimulation protects neurons against beta-amyloid toxicity. *Ann Neurol* 1997; 42: 159–163.
183. Kihara T, Shimohama S, Urushitani M, et al. Stimulation of alpha4beta2 nicotinic acetylcholine receptors inhibits beta-amyloid toxicity. *Brain Res* 1998; 792: 331–334.
184. Moon JH, Kim SY, Lee HG, et al. Activation of nicotinic acetylcholine receptor prevents the production of reactive oxygen species in fibrillar beta amyloid peptide (1–42)-stimulated microglia. *Exp Mol Med* 2008; 40: 11–18.
185. Nordberg A, Alafuzoff I, Winblad B. Nicotinic and muscarinic subtypes in the human brain: changes with aging and dementia. *J Neurosci Res* 1992; 31: 103–111.
186. Poirier J, Delisle MC, Quirion R, et al. Apolipoprotein E4 allele as a predictor of cholinergic deficits and treatment outcome in Alzheimer disease. *Proc Natl Acad Sci U S A* 1995; 92: 12 260–12 264.
187. Liu Q, Zhao B. Nicotine attenuates beta-amyloid peptide-induced neurotoxicity, free radical and calcium accumulation in hippocampal neuronal cultures. *Br J Pharmacol* 2004; 141: 746–754.
188. Garrido R, Malecki A, Hennig B, Toborek M. Nicotine attenuates arachidonic acid-induced neurotoxicity in cultured spinal cord neurons. *Brain Res* 2000; 861: 59–68.
189. Garrido R, Mattson MP, Hennig B, Toborek M. Nicotine protects against arachidonic-acid-induced caspase activation, cytochrome c release and apoptosis of cultured spinal cord neurons. *J Neurochem* 2001; 76: 1395–1403.
190. Dajas-Bailador FA, Lima PA, Wonnacott S. The alpha7 nicotinic acetylcholine receptor subtype mediates nicotine protection against NMDA excitotoxicity in primary hippocampal cultures through a Ca(2+) dependent mechanism. *Neuropharmacology* 2000; 39: 2799–2807.
191. Zeng H, Zhang Y, Peng L, et al. Nicotine and amyloid formation. *Biol Psychiatry* 2001; 49: 248–257.
192. Nordberg A, Hellstrom-Lindahl E, Lee M, et al. Chronic nicotine treatment reduces beta-amyloidosis in the brain of a mouse model of Alzheimer's disease (APPsw). *J Neurochem* 2002; 81: 655–658.
193. Hellstrom-Lindahl E, Court J, Keverne J, et al. Nicotine reduces A beta in the brain and cerebral vessels of APPsw mice. *Eur J Neurosci* 2004; 19: 2703–2710.
194. Unger C, Svedberg MM, Yu WF, et al. Effect of subchronic treatment of memantine, galantamine, and nicotine in the brain of Tg2576 (APPswe) transgenic mice. *J Pharmacol Exp Ther* 2006; 317: 30–36.

195. Shim SB, Lee SH, Chae KR, et al. Nicotine leads to improvements in behavioral impairment and an increase in the nicotine acetylcholine receptor in transgenic mice. *Neurochem Res* 2008; 33: 1783–1788.
196. Oddo S, Caccamo A, Green KN, et al. Chronic nicotine administration exacerbates tau pathology in a transgenic model of Alzheimer's disease. *Proc Natl Acad Sci U S A* 2005; 102: 3046–3051.
197. Sabbagh MN, Walker DG, Reid RT, et al. Absence of effect of chronic nicotine administration on amyloid beta peptide levels in transgenic mice overexpressing mutated human APP (Sw, Ind). *Neurosci Lett* 2008; 448: 217–220.
198. Wardlaw JM, Sandercock PA, Dennis MS, Starr J. Is breakdown of the blood-brain barrier responsible for lacunar stroke, leukoaraiosis, and dementia?. *Stroke* 2003; 34: 806–812.
199. Ballabh P, Braun A, Nedergaard M. The blood-brain barrier: an overview: structure, regulation, and clinical implications. *Neurobiol Dis* 2004; 16: 1–13.
200. Zlokovic BV. The blood-brain barrier in health and chronic neurodegenerative disorders. *Neuron* 2008; 57: 178–201.
201. Ueno M, Sakamoto H, Liao YJ, et al. Blood-brain barrier disruption in the hypothalamus of young adult spontaneously hypertensive rats. *Histochem Cell Biol* 2004; 122: 131–137.
202. Methia N, Andre P, Hafezi-Moghadam A, et al. ApoE deficiency compromises the blood brain barrier especially after injury. *Mol Med* 2001; 7: 810–815.
203. Zlokovic BV. Clearing amyloid through the blood-brain barrier. *J Neurochem* 2004; 89: 807–811.
204. Yan SD, Chen X, Fu J, et al. RAGE and amyloid-beta peptide neurotoxicity in Alzheimer's disease. *Nature* 1996; 382: 685–691.
205. Yan SD, Zhu H, Zhu A, et al. Receptor-dependent cell stress and amyloid accumulation in systemic amyloidosis. *Nat Med* 2000; 6: 643–651.
206. Donahue JE, Flaherty SL, Johanson CE, et al. RAGE, LRP-1, and amyloid-beta protein in Alzheimer's disease. *Acta Neuropathol* 2006; 112: 405–415.
207. Deane R, Du Yan S, Submamaryan RK, et al. RAGE mediates amyloid-beta peptide transport across the blood-brain barrier and accumulation in brain. *Nat Med* 2003; 9: 907–913.
208. Arancio O, Zhang HP, Chen X, et al. RAGE potentiates Abeta-induced perturbation of neuronal function in transgenic mice. *EMBO J* 2004; 23: 4096–4105.
209. Shibata M, Yamada S, Kumar SR, et al. Clearance of Alzheimer's amyloid-ss(1–40) peptide from brain by LDL receptor-related protein-1 at the blood-brain barrier. *J Clin Invest* 2000; 106: 1489–1499.
210. Kang DE, Pietrzik CU, Baum L, et al. Modulation of amyloid beta-protein clearance and Alzheimer's disease susceptibility by the LDL receptor-related protein pathway. *J Clin Invest* 2000; 106: 1159–1166.
211. Van Uden E, Mallory M, Veinbergs I, et al. Increased extracellular amyloid deposition and neurodegeneration in human amyloid precursor protein transgenic mice deficient in receptor-associated protein. *J Neurosci* 2002; 22: 9298–9304.
212. Ujiie M, Dickstein DL, Carlow DA, Jefferies WA. Blood-brain barrier permeability precedes senile plaque formation in an Alzheimer disease model. *Microcirculation* 2003; 10: 463–470.
213. Pfeifer M, Boncristiano S, Bondolfi L, et al. Cerebral hemorrhage after passive anti-Abeta immunotherapy. *Science* 2002; 298: 1379.
214. Racke MM, Boone LI, Hepburn DL, et al. Exacerbation of cerebral amyloid angiopathy-associated microhemorrhage in amyloid precursor protein transgenic mice by immunotherapy is dependent on antibody recognition of deposited forms of amyloid beta. *J Neurosci* 2005; 25: 629–636.
215. Dickstein DL, Biron KE, Ujiie M, et al. Abeta peptide immunization restores blood-brain barrier integrity in Alzheimer disease. *FASEB J* 2006; 20: 426–433.
216. Bowman GL, Kaye JA, Moore M, et al. Blood-brain barrier impairment in Alzheimer disease: stability and functional significance. *Neurology* 2007; 68: 1809–1814.
217. Farrall AJ, Wardlaw JM. Blood-brain barrier: ageing and microvascular disease - systematic review and meta-analysis. *Neurobiol Aging* 2009; 30: 337–352.
218. Mariani E, Polidori MC, Cherubini A, Mecocci P. Oxidative stress in brain aging, neurodegenerative and vascular diseases: an overview. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005; 827: 65–75.
219. Shimohama S, Tanino H, Kawakami N, et al. Activation of NADPH oxidase in Alzheimer's disease brains. *Biochem Biophys Res Commun* 2000; 273: 5–9.
220. Miller AA, Drummond GR, Schmidt HH, Sobey CG. NADPH oxidase activity and function are profoundly greater in cerebral versus systemic arteries. *Circ Res* 2005; 97: 1055–1062.
221. Park L, Anrather J, Zhou P, et al. NADPH-oxidase-derived reactive oxygen species mediate the cerebrovascular dysfunction induced by the amyloid beta peptide. *J Neurosci* 2005; 25: 1769–1777.
222. Park L, Anrather J, Girouard H, et al. Nox2-derived reactive oxygen species mediate neurovascular dysregulation in the aging mouse brain. *J Cereb Blood Flow Metab* 2007; 27: 1908–1918.
223. Park L, Zhou P, Pitstick R, et al. Nox2-derived radicals contribute to neurovascular and behavioral dysfunction in mice overexpressing the amyloid precursor protein. *Proc Natl Acad Sci U S A* 2008; 105: 1347–1352.
224. Hamel E, Nicolakakis N, Aboulkassim T, et al. Oxidative stress and cerebrovascular dysfunction in mouse models of Alzheimer's disease. *Exp Physiol* 2008; 93: 116–120.
225. Tong XK, Nicolakakis N, Kocharyan A, Hamel E. Vascular remodeling versus amyloid beta-induced oxidative stress in the cerebrovascular dysfunctions associated with Alzheimer's disease. *J Neurosci* 2005; 25: 11 165–11 174.
226. Iadecola C, Zhang F, Niwa K, et al. SOD1 rescues cerebral endothelial dysfunction in mice overexpressing amyloid precursor protein. *Nat Neurosci* 1999; 2: 157–161.
227. Varadarajan S, Yatin S, Aksenova M, Butterfield DA. Review: Alzheimer's amyloid beta-peptide-associated free radical oxidative stress and neurotoxicity. *J Struct Biol* 2000; 130: 184–208.
228. Thomas T, Thomas G, McLendon C, et al. beta-Amyloid-mediated vasoactivity and vascular endothelial damage. *Nature* 1996; 380: 168–171.
229. Folin M, Baiguera S, Conconi MT, et al. Apolipoprotein E as vascular risk factor in

neurodegenerative dementia. *Int J Mol Med* 2004; 14: 609–613.

230. Park L, Anrather J, Forster C, et al. Abeta-induced vascular oxidative stress and attenuation of functional hyperemia in mouse somatosensory cortex. *J Cereb Blood Flow Metab* 2004; 24: 334–342.
231. Nicolakakis N, Aboulkassim T, Ongali B, et al. Complete rescue of cerebrovascular function in aged Alzheimer's disease transgenic mice by antioxidants and pioglitazone, a peroxisome proliferator-activated receptor gamma agonist. *J Neurosci* 2008; 28: 9287–9296.
232. Applegate WB, Pressel S, Wittes J, et al. Impact of the treatment of isolated systolic hypertension on behavioral variables. Results from the systolic hypertension in the elderly program. *Arch Intern Med* 1994; 154: 2154–2160.
233. Prince MJ, Bird AS, Blizard RA, Mann AH. Is the cognitive function of older patients affected by antihypertensive treatment? Results from 54 months of the Medical Research Council's trial of hypertension in older adults. *BMJ* 1996; 312: 801–805.
234. Cervilla JA, Prince M, Joels S, et al. Long-term predictors of cognitive outcome in a cohort of older people with hypertension. *Br J Psychiatry* 2000; 177: 66–71.
235. Lithell H, Hansson L, Skoog I, et al. The Study on COgnition and Prognosis in the Elderly (SCOPE); outcomes in patients not receiving add-on therapy after randomization. *J Hypertens* 2004; 22: 1605–1612.
236. Trenkwalder P. The Study on COgnition and Prognosis in the Elderly (SCOPE)–recent analyses. *J Hypertens Suppl* 2006; 24: S107–S114.
237. Tzourio C, Anderson C, Chapman N, et al. Effects of blood pressure lowering with perindopril and indapamide therapy on dementia and cognitive decline in patients with cerebrovascular disease. *Arch Intern Med* 2003; 163: 1069–1075.
238. Staessen JA, Fagard R, Thijs L, et al. Randomised double-blind comparison of placebo and active treatment for older patients with isolated systolic hypertension. The Systolic Hypertension in Europe (Syst-Eur) Trial Investigators. *Lancet* 1997; 350: 757–764.
239. Forette F, Seux ML, Staessen JA, et al. Prevention of dementia in randomised double-blind placebo-controlled Systolic Hypertension in Europe (Syst-Eur) trial. *Lancet* 1998; 352: 1347–1351.
240. Forette F, Seux ML, Staessen JA, et al. The prevention of dementia with antihypertensive treatment: new evidence from the Systolic Hypertension in Europe (Syst-Eur) study. *Arch Intern Med* 2002; 162: 2046–2052.
241. Poon IO. Effects of antihypertensive drug treatment on the risk of dementia and cognitive impairment. *Pharmacotherapy* 2008; 28: 366–375.
242. Nilsson SE, Read S, Berg S, et al. Low systolic blood pressure is associated with impaired cognitive function in the oldest old: longitudinal observations in a population-based sample 80 years and older. *Aging Clin Exp Res* 2007; 19: 41–47.

SPECIAL FEATURE

First Season of 2009 H1N1 Influenza

Kristin A. Swedish, MD, Gina Conenello, and Stephanie H. Factor, MD, MPH

Mount Sinai School of Medicine, New York, NY

OUTLINE

Virology of Influenza
Historical Pandemics
Swine-origin Influenzas Prior to 2009
2009 H1N1 Influenza Pandemic
Clinical Disease
Diagnosis and Treatment
Epidemiology
Prevention
Conclusion

ABSTRACT

The 2009 H1N1 influenza virus is responsible for the first influenza pandemic of the 21st century. Within 5 months after its emergence in Mexico, it had affected more than 290,000 patients and caused at least 3486 deaths on 6 continents. This report provides information about influenza viruses, previous pandemics, and the virology, clinical features, and epidemiology of 2009 H1N1 influenza in order to elucidate the current pandemic and outline the questions that remain. ***Mt Sinai J Med 77:103–113, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: H1N1, influenza, pandemic, swine flu.

Outbreaks of respiratory illnesses and an increased number of influenza-like illnesses were reported in Mexico throughout March 2009.[1] On April 21, 2009, the Centers for Disease Control and Prevention (CDC) reported that 2 Americans were infected with a novel influenza A virus (H1N1) that was likely related.[2] On April 26, 2009, after 20 Americans were infected, the presence of the novel influenza virus was declared an emergency by the US Department of Health and Human Services.[3] On June 11, 2009, the World Health Organization declared 2009 H1N1 influenza a phase 6 (global) pandemic after 74 countries had reported disease (Figure 1).[4] This report provides the background for the 2009 H1N1 influenza virus so that we can better understand the current pandemic and outline the questions that remain.

Address Correspondence to:

Kristin A. Swedish
Mount Sinai School of Medicine
New York, NY
Email: kristin.swedish@mssm.edu

VIROLOGY OF INFLUENZA

Influenza A, from the family *Orthomyxoviridae*, is a single-stranded, enveloped RNA virus. The viral particle has 2 types of glycoprotein spikes that project from its surface: hemagglutinin (HA) and neuraminidase (NA). HA recognizes and binds to cell surface sialic acids, compounds found on the cells of many animal species. This leads to the fusion of the viral and cell membranes and allows the virus to enter the cell. After intracellular viral RNA transcription, replication, and translation in the host cell, NA enables newly formed viruses to leave by cleaving the HA–sialic acid bond.[6,7]

The surface glycoprotein spikes are 2 of the 11 proteins encoded in the 8-segment influenza A RNA viral genome (Table 1). HA is encoded by segment 4; NA is encoded by segment 6. Polymerase basic 1 frame 2 (PB1-F2) is a nonstructural protein encoded by segment 2 that may or may not be present in the virus. When present, it is a proapoptotic protein causing programmed cell death and increased virulence.[6]

The human immune system specifically responds to the HA and NA glycoproteins. Influenza A viruses are subtyped according to the type of HA and NA glycoproteins that they possess. There are 16 different possible subtypes for HA and 9 for NA. Only 3 HA subtypes (H1, H2, and H3) and 2 NA subtypes (N1 and N2) have been found in human influenza viruses.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20164

© 2010 Mount Sinai School of Medicine

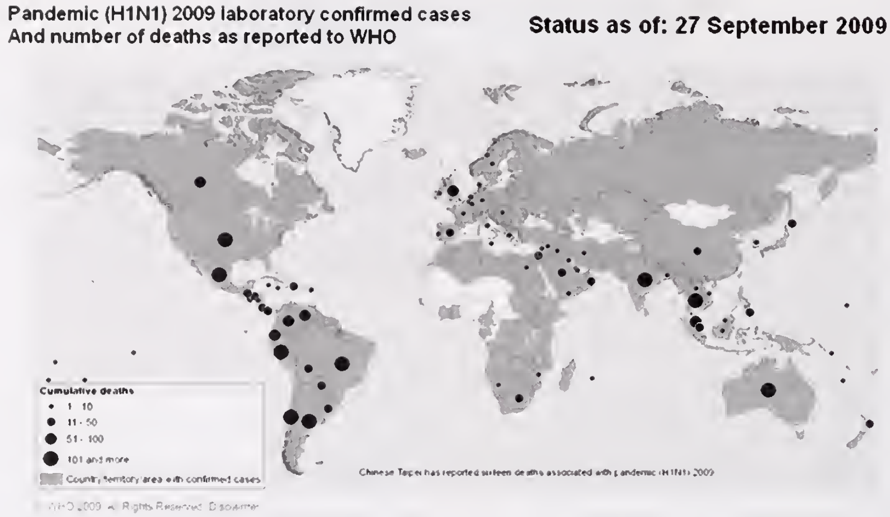

Pandemic (H1N1) 2009 laboratory confirmed cases
And number of deaths as reported to WHO
Status as of: 27 September 2009
Cumulative deaths
1 - 10
11 - 50
51 - 100
101 and more
Country/territory/area with confirmed cases
Chinese Taipei has reported sixteen deaths associated with pandemic (H1N1) 2009
© WHO 2009. All Rights Reserved. Disclaimer

Fig 1. Laboratory-confirmed cases of 2009 H1N1 influenza and number of deaths as reported to the World Health Organization (status as of September 27, 2009). Reprinted with permission from the World Health Organization.[5] Copyright 2009, World Health Organization.

Table 1. *Influenza A Viral Genome.*

Segment	Protein	Function
1	PB2	Messenger RNA cap recognition during RNA transcription
2	PB1	Polymerase activity
	PB1-F2	Accessory protein; proapoptotic
3	PA	Endonuclease activity; necessary for RNA elongation
4	HA	Receptor binding and fusion
5	NP	RNA binding protein; regulates nuclear viral import
6	NA	Virus release
7	M1	Matrix protein; regulates viral RNA export from host cell nucleus
	M2	Forms ion channel; assists with viral uncoating and assembly
8	NS1	Regulates host gene expression; interferon antagonist
	NS2	Nuclear export of RNA

NOTE: This table was adapted from Bouvier and Palese[6] and Dias *et al.*[8]

Abbreviations: HA, hemagglutinin; M1, matrix 1; M2, matrix 2; NA, neuraminidase; NP, nucleoprotein; NS1, nonstructural 1; NS2, nonstructural 2; PA, polymerase acidic; PB1, polymerase basic 1; PB1-F2, polymerase basic 1 frame 2; PB2, polymerase basic 2.

Changes in these glycoproteins create new strains of influenza A. After recovering from an influenza infection, humans are still susceptible to subsequent influenza infections caused by viruses with different glycoprotein subtypes.[6,7]

Changes in the glycoproteins occur via 2 main mechanisms. Antigenic drift occurs when errors in the viral genome occur during replication and cause small changes in the viral proteins. Antigenic shift occurs when 2 different influenza A subtypes exchange full RNA segments. Antigenic shift is possible when a cell is coinfected with multiple influenza A subtypes.[6,9] It is currently believed that all known pandemic influenzas have been caused by antigenic shift resulting in novel influenza viruses to which humans are both susceptible and immunologically naïve.[6,10,11]

> *It is currently believed that all known pandemic influenzas have been caused by antigenic shift resulting in novel influenza viruses to which humans are both susceptible and immunologically naïve.*

Genetic analysis has demonstrated that all mammalian influenza viruses are originally derived from avian influenza viruses. Aquatic birds are the asymptomatic primary natural reservoir for all subtypes of influenza A. Influenza viruses replicate predominantly in avian intestinal tract cells and are excreted in high concentrations in feces. These viruses of avian origin are capable of infecting mammals, including seals, whales, horses, dogs, cats, pigs, and humans. Over time, viruses have evolved into host-specific lineages in horses, humans, and pigs.[12]

Species have specific sialic acid configurations on their cell surfaces that are recognized by species-specific influenza viruses. Influenza viruses of avian origin preferentially recognize the avian-specific sialic acid configuration found predominantly on bird intestinal tract cells. Influenza viruses of human origin preferentially recognize the human-specific sialic acid configuration found on human respiratory tract cells. Both avian-type and human-type sialic acid configurations are found on swine respiratory tract cells. Swine respiratory tract cells, therefore, may be coinfected with avian and human influenza viruses in addition to influenza viruses also circulating in the pig population. Once coinfected, swine cells can act as mixing vessels and enable the reassortment of viral segments from these 3 different species, which causes antigenic shift.[6,10,11]

HISTORICAL PANDEMICS

The 1918 influenza pandemic (Spanish influenza) was caused by an H1N1 influenza virus. The first influenza cases occurred in March 1918 in both Europe and the United States with a high attack rate and a low case fatality rate (a rate of 1.1% was estimated in US Army camps[13]). A second wave spread worldwide in the fall of 1918, and a third wave affected many countries in early 1919. The second and third waves were marked by a higher case fatality rate (a rate of 4.7% was estimated in US Army camps[13]). Approximately 40 million people died worldwide between 1918 and 1919 for an overall case fatality rate of over 2.5%. Nearly 50% of the influenza-related deaths occurred among young adults (20–40 years old), and less than 1% occurred among adults over the age of 65 years. It is theorized that older adults had partial protection from exposure to a similar virus that had been circulating before 1889.[14]

In 1918–1919, swine were also infected with a respiratory illness that resembled human influenza, and this suggested that the 1918 H1N1 virus was able to infect pigs. Scientists believe that the 1918 H1N1 virus continued to circulate in the swine population for many decades after this introduction.[15] In 1931, Robert Shope, a veterinarian, demonstrated that human serum from adults who survived the 1918 pandemic was protective against infection with a swine influenza virus isolated in 1930, and this suggested a close antigenic similarity between the swine and human viruses.[16]

Three subsequent influenza pandemics were also caused by antigenic shift. In 1957, H2N2 emerged and caused many deaths in the immunologically naïve population. Approximately 70,000 people died in the United States alone. In 1968, H3N2 emerged but had lower morbidity and mortality rates because of its antigenic similarity to the 1957 H2N2 strain: it differed in only 2 of its 11 viral proteins. In 1977, another H1N1 strain emerged, with a lower attack rate among people born before 1950, who had likely been exposed and developed protective immunity to earlier circulating H1N1 viruses.[7]

Since 1977, H1N1 viruses have been circulating in the human population and causing seasonal influenza. For approximately the past 10 years, 3 influenza viruses have contributed to seasonal influenza: 2 influenza A viruses (H1N1 and H3N2) and 1 influenza B virus. These 3 viruses circulate between November and March in the Northern Hemisphere and between April and September in the Southern Hemisphere.[6] Typical seasonal influenza causes 250,000 to 500,000 deaths annually,[17] approximately 35,000 of which are in the United States. Approximately 200,000 people are hospitalized annually in the United States with seasonal flu.[18]

SWINE-ORIGIN INFLUENZAS PRIOR TO 2009

Although swine-origin influenza A viruses typically do not infect humans, occasional infections have occurred in recent years, usually after direct exposure to pigs. The severity of swine-origin influenza infection in humans ranges from mild to severe. The first "swine flu" scare occurred at the Fort Dix Army base in New Jersey in 1976, when approximately 200 people were infected with a novel H1N1 swine influenza virus that caused several serious infections and 1 death. Fear of a pandemic prompted a mass immunization campaign. Despite the panic that it caused, the swine influenza pandemic never materialized.[19] The vaccination campaign was stopped after 532 cases of Guillain-Barré syndrome resulted in 32 deaths.[15]

Since 1976, there have been scattered reports of swine influenza. According to the CDC, 1 swine influenza infection was reported in the United States every 1 to 2 years. However, between December 2005 and February 2009, 11 nonfatal cases of swine influenza infection were reported. The first case in December 2005 was identified as the first human infection with a "triple-reassortment swine influenza," a combination of avian, human, and swine viral gene segments. As demonstrated in Figure 2, the proteins polymerase basic 2 (PB2) and polymerase acidic (PA; gene segments 1 and 3) were of

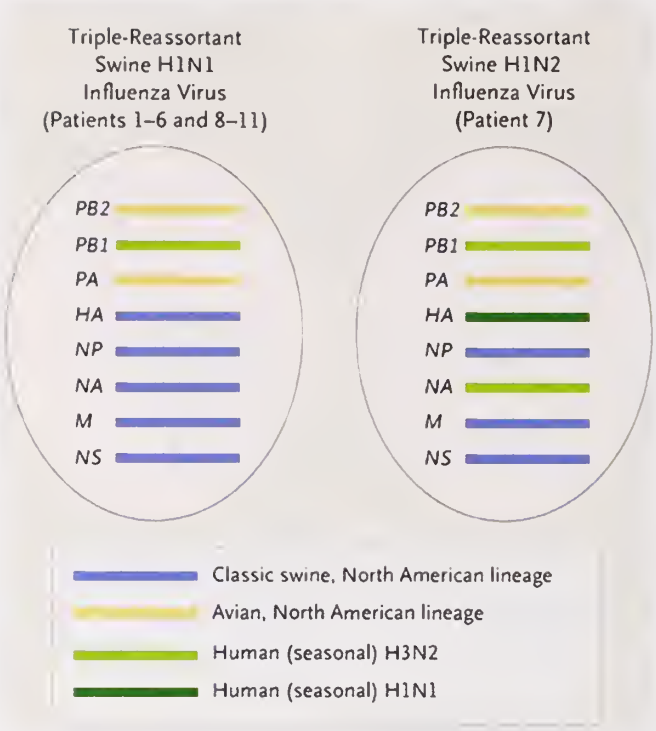

Triple-Reassortant Swine H1N1 Influenza Virus (Patients 1–6 and 8–11)
Triple-Reassortant Swine H1N2 Influenza Virus (Patient 7)
PB2
PB1
PA
HA
NP
NA
M
NS
Classic swine, North American lineage
Avian, North American lineage
Human (seasonal) H3N2
Human (seasonal) H1N1

Fig 2. Triple-reassortant swine influenza A isolated from 11 patients between December 2005 and February 2009 in the United States. It is common practice in the virology community to call these proteins by their abbreviations without describing what they stand for. Since we have done it once in text and in the previous Figure, we need to keep explaining the abbreviations. Reprinted with permission from the *New England Journal of Medicine.*[20] Copyright 2009, Massachusetts Medical Society.

North American avian origin. The protein polymerase basic 1 (PB1; segment 2) was from human seasonal H3N2. The HA, nucleoprotein (NP), NA, matrix (M), and nonstructural (NS) proteins (segments 4–8) were of North American classic swine origin. All subsequent human infections of swine origin had this triple reassortment.[20]

2009 H1N1 INFLUENZA PANDEMIC

The first confirmed case patient of 2009 H1N1 influenza in Mexico was a child in La Gloria, a village that is home to many industrial hog farms. Direct swine-to-human transmission in Mexico is the presumed source of the pandemic. The first 2 confirmed case patients of 2009 H1N1 influenza in the United States were children in Southern California who had not had any contact with pigs, and this suggested human-to-human transmission. Indeed, many of the initial American cases could not be linked to travel to Mexico or to exposure to another confirmed case patient.[21]

The 2009 H1N1 influenza virus was identified in April 2009 as a previously unknown H1N1 virus. As demonstrated in Figure 3, the proteins PB2, PB1, PA, NA, NP, and NS (gene segments 1–5 and 8) were similar to those previously found in triple-reassortment swine influenza viruses. The remaining 2 proteins, NA and M (segments 6 and 7), were from the Eurasian swine genetic lineage first noted in 1979 [the proteins HA, NP, and NS (gene segments 4, 5, and 8) had been in the classic swine lineage prior to triple reassortment[22]].

Despite the novel genomic reassortment of the 2009 H1N1 influenza virus, 3 components of the virus's genome suggest that morbidity and mortality rates will be lower than originally feared. First, unlike the 1918, 1957, and 1968 viruses, the 2009 H1N1 influenza virus does not contain the proapoptotic PB1-F2 accessory protein.[19] Laboratory studies have shown that expression of PB1-F2 enhances inflammation during the primary viral infection of mice and increases both the frequency and severity of secondary bacterial pneumonia.[23] Second, although the 2009 H1N1 influenza virus

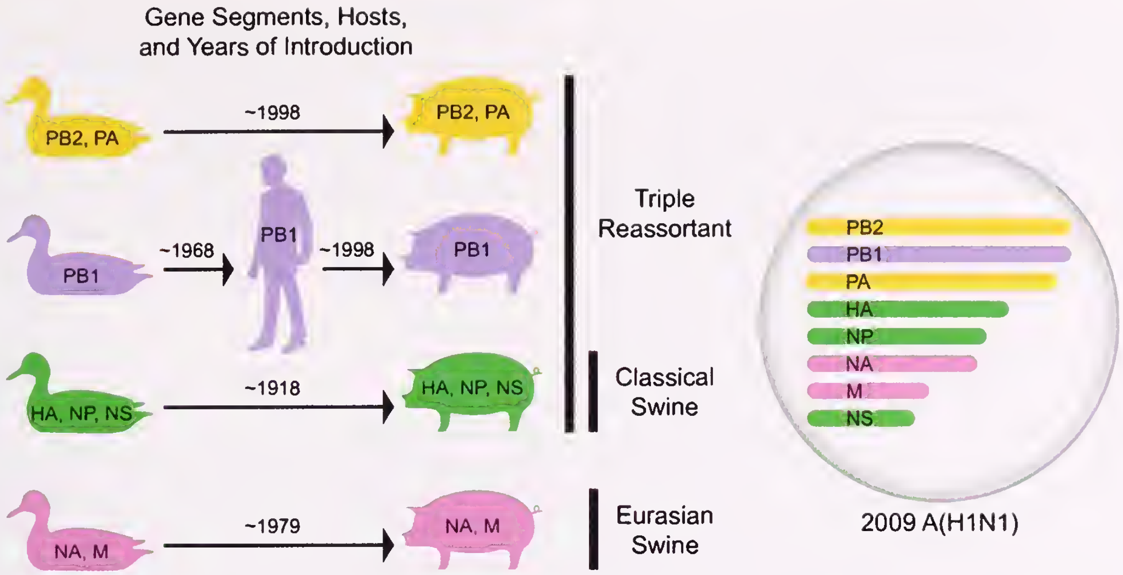


Fig 3. Host and lineage origins for the gene segments of the 2009 H1N1 influenza virus. Reprinted with permission from *Science*.[22] Copyright 2006, American Association for the Advancement of Science.

Despite the novel genomic reassortment of the 2009 H1N1 influenza virus, 3 components of the virus's genome suggest that morbidity and mortality rates will be lower than originally feared.

has demonstrated resistance to the adamantane M2 inhibitor class of influenza drugs (amantadine and rimantadine), it is sensitive to the NA inhibitors (oseltamivir and zanamivir).[19] Finally, its genome is sufficiently similar to those of previously circulating influenza viruses to allow a degree of protective immunity in portions of the human population. In microneutralization assays performed by the CDC, 33% of adults over 60 years demonstrated cross-reactive antibody levels associated with a 50% reduction in the risk for influenza infection.[24]

CLINICAL DISEASE

Clinically, the signs and symptoms of 2009 H1N1 influenza are similar to those of seasonal influenza. Patients have a febrile respiratory illness with cough, sore throat, myalgias, headache, chills, and/or fatigue. Case series have shown that 25%[25] to 40%[26] of confirmed case patients have also had diarrhea or vomiting, unlike seasonal influenza. The incubation period ranges from 2 to 7 days. Its infectious period, when viral shedding occurs, is presumed to be similar to that of seasonal flu. It ranges from 1 day before symptom onset to at least 5 to 7 days after symptom onset.[25] The majority of 2009 H1N1 influenza infections have been self-limited and mild in nature.[27] Some cases of severe illness, causing respiratory failure requiring intensive care, have been reported.[28]

The majority of 2009 H1N1 influenza infections have been self-limited and mild in nature.

DIAGNOSIS AND TREATMENT

The CDC recommends 2009 H1N1 influenza laboratory testing for hospitalized patients with suspected influenza; patients for whom an influenza diagnosis will affect decision making regarding clinical care, infection control, or management of close contacts; and patients who died of suspected influenza.[29] Antiviral treatment is recommended for all hospitalized patients with confirmed or suspected 2009 H1N1 influenza and those at high risk for complications. Treatment with zanamivir or oseltamivir is best when initiated within 48 hours of illness onset. The treatment duration is 5 days. Some studies of hospitalized patients have shown a benefit from antiviral treatment when treatment was initiated more than 48 hours after the onset of illness. Therefore, patients at high risk for complications may initiate treatment more than 48 hours after illness onset. Antiviral prophylaxis with a 10-day course of oseltamivir or zanamivir is recommended for anyone at high risk for complications who has had close contact with cases of confirmed, probable, or suspected 2009 H1N1 influenza or seasonal influenza. The CDC also recommends prophylaxis for health care workers who have had unprotected close-contact exposure to a person with confirmed, probable, or suspected 2009 H1N1 influenza or seasonal influenza during that person's infectious period. Antiviral prophylaxis of healthcare workers should be initiated within 48 hours of contact with an infectious person.[30]

There have been scattered reports of patients with oseltamivir-resistant strains of 2009 H1N1 influenza. Two unrelated immunosuppressed patients, both in Seattle, developed resistance to oseltamivir during the course of treatment for 2009 H1N1 influenza.[31] Oseltamivir-resistant 2009 H1N1 influenza was also reported in 2 adolescents who had received oseltamivir in a mass chemoprophylaxis program during a summer camp outbreak in North Carolina.[32] In Hong Kong, a previously healthy 16-year-old girl who had not received oseltamivir prophylactically or for treatment tested positive for oseltamivir-resistant 2009 H1N1 influenza.[33]

EPIDEMIOLOGY

Several epidemiological factors are unique to the 2009 H1N1 influenza pandemic. The typical influenza season in the Northern Hemisphere ends in March, yet 2009 H1N1 influenza cases were reported throughout the spring and summer and into the fall of 2009. This extension of the traditional influenza season in the Northern Hemisphere is unusual. The typical influenza season in the Southern Hemisphere ends in September, yet 2009 H1N1 influenza cases were reported through early October.[34]

Second, the age distribution of morbidity and mortality due to 2009 H1N1 influenza is different from that due to seasonal influenza. Although seasonal influenza predominantly affects children under 5 years of age and adults over 64 years of age, the majority of morbidity and mortality from 2009 H1N1 influenza has occurred in persons younger than 45 years. Country-specific and state-specific data from the first season showed that for all laboratory-confirmed cases the median age of persons with clinical disease ranged from 12 to 19 years, the median age of persons hospitalized with 2009 H1N1 influenza ranged from 16 to 27.5 years, the median age of persons hospitalized with complications from 2009 H1N1 influenza ranged from 38 to 45 years, and the median age of persons deceased because of 2009 H1N1 influenza ranged from 26 to 45 years (Table 2).

Although seasonal influenza predominantly affects children under 5 years of age and adults over 64 years of age, the majority of morbidity and mortality from 2009 H1N1 influenza has occurred in persons younger than 45 years.

Epidemiological data collected during the first season have demonstrated that certain groups are at higher risk for complications of 2009 H1N1 influenza. These groups, which are similar to those who are at increased risk for complications of seasonal influenza, include children under the age of 2 years, pregnant women and women up to 2 weeks postpartum (including women who have suffered pregnancy loss), children and adults with chronic diseases (including pulmonary, cardiovascular, renal, hepatic, hematological, and metabolic disorders), children and adults with disorders that compromise respiratory function or increase the risk for aspiration, and children and adults who are immunosuppressed (including immunosuppression caused by medications or human immunodeficiency virus). Unlike seasonal influenza, adults over 65 years appear to have age-related protective immunity to 2009 H1N1 influenza. However, when they acquire 2009 H1N1 influenza, adults over 65 years are at increased risk for complications, and this mandates their inclusion in the list of groups at higher risk for complications.[30]

Of patients hospitalized with 2009 H1N1 influenza during the first season, the proportion with underlying medical conditions ranged from 30%[38] to 76%.[42] The most common underlying medical conditions were chronic lung disease, including asthma and chronic obstructive pulmonary disease, immunosuppression, chronic cardiac disease, diabetes, and obesity. The same medical conditions were common among patients who died of 2009 H1N1 influenza. The proportion of patients with underlying medical conditions who died ranged from 33%[38] to 71%.[39]

Like seasonal influenza, pregnancy appears to be an important risk factor for severe disease with

Table 2. *State-Specific and Country-Specific Epidemiological Data for Confirmed Cases of 2009 H1N1 Influenza.*

Authors	Date of Publication	Location	Study Period	Setting	Number Affected	Age Affected (Median, Range)	CFR	Age at Death (Median, Range)	% with Underlying Medical Conditions	Underlying Medical Conditions
Ciblak *et al.*[35]	August 13, 2009	Turkey	May 15 to July 17, 2009	Confirmed H1N1 cases	128	52% between 19–34 years	0	N/A	–*	–*
CDC[36]	May 22, 2009	California	April 15 to May 17, 2009	First 30 patients hospitalized with confirmed H1N1	30	27.5 years, 27 days to 89 years	0	N/A	64%	Chronic lung disease (37%) Immunosuppression (20%) Chronic cardiac disease (17%) Diabetes (13%) Obesity (13%)
CDC[37]	June 5, 2009	Mexico	March 1 to May 29, 2009	Laboratory-confirmed H1N1 cases	5337	41.9% <15 years 32.3% between 15–29 years	–*	55.7% between 30 to 59 years	–*	–*
CDC[38]	July 17, 2009	Michigan	May 26 to June 18, 2009	10 ICU patients with H1N1 and ARDS	10	45 years, 21 to 53 years	30%	43 years, 28 to 44 years	30% (hospitalized) 33% (died)	Asthma (20% hospitalized: 33% died)
CDC[39]	August 28, 2009	Chicago, IL	April 24 to July 25, 2009	Laboratory-confirmed H1N1 cases and hospitalized patients with confirmed H1N1	1557	12 year, 24 days to 91 years	0.45%	32 years, 20 to 54 years	71% (died)	Asthma (21% hospitalized; 14% died)
					205	16 year, 24 days to 91 years				Immunosuppression (29% died) Pregnancy (7% hospitalized; 14% died)
Gilsdorf and Poggensee[40]	August 27, 2009	Germany	April 20 to August 10, 2009	Confirmed H1N1 cases	9950	19% between 0–89 years 77% between 10–29 years	–*	–*	3.3%	Respiratory disease (45%) Cardiovascular disease (15%) Diabetes (9%) Obesity (6%) Pregnancy (5%)
Health Protection Agency *et al.*[41]	June 4, 2009	United Kingdom	April 27 to May 31, 2009	Confirmed H1N1 cases	252	12 years, 0 to 73 years	0	N/A	–*	–*
New York City Department of Health and Mental Hygiene[42]	July 8, 2009	New York, NY	April 25 to July 8, 2009	Hospitalized patients with confirmed H1N1	909	21 years	4.95%	44 years, 2 months to 83 years	76% (hospitalized)	Asthma (30% hospitalized; 29% died)
						43% <18 years, 77% <50 years			70% (died)	Diabetes (13% hospitalized; 31% died)

Table 2. *(Continued).*

Authors	Date of Publication	Location	Study Period	Setting	Number Affected	Age Affected (Median, Range)	CFR	Age at Death	% with Underlying Medical Conditions	Underlying Medical Conditions
Jain *et al.*[26]	October 8, 2009	United States	May 1 to June 9, 2009	First 272 patients hospitalized with confirmed H1N1	272	21 years, 21 days to 86 years	6.99%	26 years, 1.3 to 57 years	73% (hospitalized) 68% (died)	Heart disease (12% hospitalized; 18% died) Immune deficiency (9% hospitalized; 22% died) Asthma or COPD (36% hospitalized; 16% died) Diabetes (15% hospitalized) Immunosuppression (15% hospitalized) Neurological disease (7% hospitalized; 21% died) Pregnancy (7% hospitalized; 16% died)
Munayco *et al.*[43]	August 13, 2009	Peru	May 9 to July 17, 2009	Confirmed H1N1 cases	1771	13 years, 0–87 years	0.33%	–*	–*	–*
Perez-Padilla *et al.*[44]	June 29, 2009 (online)	Mexico	March 24 to April 24, 2009	Hospitalized patients with confirmed H1N1 and pneumonia	18	38 years, 9 months to 61 years	39%	45 years, 9 to 52 years	44.4% (hospitalized) 42.9% (died)	Hypertension (17% hospitalized; 14% died) Diabetes (17% hospitalized) Asthma (11% hospitalized) Obstructive sleep apnea (6% hospitalized; 14% died)
Vaillant etal.[45]	August 20, 2009	Worldwide	April 27 to July 16, 2009	First 684 deaths of patients with confirmed H1N1	126,168	N/A	0.6%	37 years, 0 to 87 years	53% (died)	Diabetes and/or obesity (28%) Respiratory disease (19%) Heart disease (19%)

Abbreviations: ARDS, acute respiratory distress syndrome; CDC, Centers for Disease Control and Prevention; CFR, case fatality rate; COPD, chronic obstructive pulmonary disease; ICU, intensive care unit; N/A, not applicable.
*Not discussed by the authors.

Table 3. *Laboratory-Confirmed Cases of 2009 H1N1 Influenza Reported to the World Health Organization as of July 6, 2009.*

Country	Cases	Deaths	CFR (%)
Argentina	2,485	60	2.41
Uruguay	195	4	2.05
Dominican Republic	108	2	1.85
Colombia	118	2	1.69
Mexico	10,262	119	1.16
Costa Rica	277	3	1.08
Paraguay	106	1	0.94
Honduras	123	1	0.81
Guatemala	286	2	0.70
United States	33,902	170	0.50
Thailand	2,076	7	0.34
Canada	7,983	25	0.31
New Zealand	1,059	3	0.28
Chile	7,376	14	0.19
Australia	5,298	10	0.19
Brazil	737	1	0.14
Spain	776	1	0.13
Philippines	1,709	1	0.06
United Kingdom	7,447	3	0.04

NOTE: This table was adapted from the World Health Organization Web site.[46]
Abbreviation: CFR, case fatality rate.

the acquisition of 2009 H1N1 influenza. In New York City, 16% of hospitalized women of childbearing age were pregnant.[42] Five of the first 30 patients (16.7%) hospitalized in California were pregnant, 2 of whom developed complications (spontaneous abortion and premature rupture of membranes).[36] Of the first 272 patients hospitalized in the United States, 18 (7%) were pregnant women, 6 of whom had another underlying medical condition. Nineteen of these hospitalized patients died, and 3 had been pregnant.[26] Global death data show that approximately 10% of female patients who died were pregnant or had delivered at the time of their death. Of note, the available data suggest that up to half of the women also had underlying medical conditions, including asthma, obesity, and heart disease.[45]

The case fatality rates for 2009 H1N1 influenza have varied from country to country (Table 3). The worldwide alert to the disease in Mexico likely provided public health practitioners in other countries time to implement preparedness plans.

PREVENTION

To promote widespread immunity to the virus, four 2009 H1N1 influenza vaccines have been developed and approved for use by the Food and Drug Administration. There is 1 live attenuated virus formulation that is administered intranasally and 3 inactivated virus formulations that are administered intramuscularly.[47] The CDC's Advisory Committee on Immunization Practices recommends that vaccination efforts focus initially on 5 groups of people: pregnant women, those who live with and care for children younger than 6 months, health care and emergency service personnel, people between the ages of 6 months and 24 years of age, and those between 25 and 64 years old with health conditions associated with a higher risk of complications from 2009 H1N1 influenza. Because of the decreased risk of infection for those 65 years old and older, the CDC recommends their vaccination only after the vaccination of younger people has been completed.[48]

> *The CDC recommends that vaccination efforts focus initially on 5 groups of people: pregnant women, those who live with and care for children younger than 6 months, health care and emergency service personnel, people between the ages of 6 months and 24 years of age, and those between 25 and 64 years old with health conditions associated with a higher risk of complications from 2009 H1N1 influenza.*

The effectiveness of the vaccine against 2009 H1N1 influenza had been a main concern during production. A vaccine immunogenicity trial was conducted in China with 2200 subjects ranging in age from 3 to 77 years. A hemagglutination inhibition titer of 1:40 or more within 3 weeks of vaccination was considered protective. A single 15-µg dose of the inactivated vaccine produced a protective serologic immune response in subjects ranging in age from 12 to 60 years. Children that were 3 to 11 years old and adults that were 61 years or older required a second dose to produce a similar response.[49] Preliminary results from a vaccine immunogenicity trial in American children showed that a single 15-µg dose of the inactivated vaccine produced a protective serologic immune response in children that were 10 to 17 years old, whereas children that were 6 months to 9 years old had a less robust response.[50] The CDC therefore recommends that children who are 6 months to 9 years old receive 2 doses of vaccine approximately 4 weeks apart, whereas people over the age of 10 years require only 1 dose.[47]

CONCLUSION

The activity of the 2009 H1N1 influenza virus during the 2009–2010 influenza season will hopefully answer the numerous questions that remain. The 1918 Spanish influenza pandemic was characterized by multiple waves in quick succession that occurred too quickly to have been caused by antigenic drift. The first wave had a low case fatality rate and was followed by a second wave (and, in some countries, a third wave) with a much higher case fatality rate. There is great concern that 2009 H1N1 influenza will follow a similar pattern.

Currently, the H1N1 viruses from around the world remain genetically close.[28] Of great concern is whether 2009 H1N1 influenza will mutate to become more virulent by either antigenic drift or acquisition of the proapoptotic PB2-F1 accessory protein. Sporadic cases of oseltamivir-resistant 2009 H1N1 influenza viruses have raised concern about widespread resistance.

Mass immunization campaigns have been shown to provide herd immunity against seasonal influenza. When 70% of 3- to 17-year-old children in one Moscow neighborhood were vaccinated in 2001, the incidence rate of influenza-like illness in unvaccinated adults over 60 years was one-third of the incidence rate of influenza-like illness in unvaccinated adults over 60 years in an unvaccinated neighborhood.[51] We do not yet know the vaccination coverage rate required for herd immunity against 2009 H1N1 influenza. Hopefully, herd immunity will be achieved early in the influenza season.

It is unclear how the 2009 H1N1 influenza pandemic will continue and when it will end. Global surveillance of epidemiological, clinical, and virological data over the 2009–2010 influenza season will likely answer these remaining questions and enable the medical and public health communities to remain prepared.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Centers for Disease Control and Prevention. Outbreak of swine-origin influenza A (H1N1) virus infection—Mexico, March-April 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 467–470.
2. Centers for Disease Control and Prevention. Swine influenza A (H1N1) infection in two children—Southern California, March-April 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 400–402.
3. US Department of Health and Human Services. HHS declares public health emergency for swine flu. Available at: http://www.hhs.gov/news/press/2009pres/04/20090426a.html. Accessed November 2009.
4. World Health Organization. Influenza A (H1N1)—update 23. Available at: http://www.who.int/csr/don/2009_05_09/en/index.html. Accessed November 2009.
5. World Health Organization. Pandemic (H1N1) 2009—update 66. Available at: http://www.who.int/csr/don/2009_09_18/en/index.html. Accessed November 2009.
6. Bouvier NM, Palese P. The biology of influenza viruses. *Vaccine* 2008; 26(suppl 4): D49–D53.
7. Palese P. Influenza: old and new threats. *Nat Med* 2004; 10(suppl 12): S82–S87.
8. Dias A, Bouvier D, Crépin T, et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* 2009; 458: 914–918.
9. National Institute of Allergy and Infectious Disease. Flu (influenza): what's the difference between seasonal flu, pandemic flu, and avian (bird) flu? Available at: http://www3.niaid.nih.gov/topics/Flu/understanding Flu/DefinitionsOverview.htm. Accessed November 2009.
10. This Week in Virology. Interview with Dr. Peter Palese. Available at: http://www.twiv.tv/TWiV-Special-Palese-050609.pdf. Accessed November 2009.
11. Ma W, Lager KM, Vincent AL, et al. The role of swine in the generation of novel influenza viruses. *Zoonoses Public Health* 2009; 56: 326–337.
12. Webster RG. Influenza: an emerging disease. *Emerg Infect Dis* 1998; 4: 436–441.
13. Barry JM, Viboud C, Simonsen L. Cross-protection between successive waves of the 1918–1919 influenza pandemic: epidemiological evidence from US Army camps and from Britain. *J Infect Dis* 2008; 198: 1427–1434.
14. Taubenberger JK, Morens DM. 1918 influenza: the mother of all pandemics. *Emerg Infect Dis* 2006; 12: 15–22.
15. Zimmer SM, Burke DS. Historical perspective—emergence of influenza A (H1N1) viruses. *N Engl J Med* 2009; 361: 279–285.
16. Stein RA. Lessons from outbreaks of H1N1 influenza. *Ann Intern Med* 2009; 151: 59–62.
17. World Health Organization. Fact sheet no. 211: influenza (seasonal). Available at: http://www.who.int/mediacentre/factsheets/fs211/en/index.html. Accessed November 2009.
18. Centers for Disease Control and Prevention. Seasonal influenza: the disease. Available at: http://www.cdc.gov/flu/about/disease/index.htm. Accessed November 2009.
19. Wang TT, Palese P. Unraveling the mystery of swine influenza virus. *Cell* 2009; 137: 983–985.
20. Shinde V, Bridges CB, Uyeki TM, et al. Triple-reassortant swine influenza A (H1) in humans in the United States, 2005–2009. *N Engl J Med* 2009; 360: 2616–2625.
21. Centers for Disease Control and Prevention. Update: swine influenza A (H1N1) infections—California and Texas, April 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 435–437.

DOI:10.1002/MSJ

22. Garten RJ, Davis CT, Russell CA, et al. Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. *Science* 2009; 325: 197–201.
23. McAuley JL, Hornung F, Boyd KL, et al. Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell Host Microbe* 2007; 2: 240–249.
24. Centers for Disease Control and Prevention. Serum cross-reactive antibody response to a novel influenza A (H1N1) virus after vaccination with seasonal influenza vaccine. *MMWR Morb Mortal Wkly Rep* 2009; 58: 521–524.
25. Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med* 2009; 360: 2605–2615.
26. Jain S, Kamimoto L, Bramley AM, et al. Hospitalized patients with 2009 H1N1 influenza in the United States. *N Engl J Med* 2009; 361: 1935–1944.
27. Centers for Disease Control and Prevention. Questions & answers: 2009 H1N1 flu (swine flu) and you. Available at: http://www.cdc.gov/h1n1flu/qa.htm. Accessed November 2009.
28. World Health Organization. Preparing for the second wave: lessons from current outbreaks. Available at: http://www.who.int/csr/disease/swineflu/notes/h1n1_second_wave_20090828/en/index.html. Accessed November 2009.
29. Centers for Disease Control and Prevention. Interim recommendations for clinical use of influenza diagnostic tests during the 2009–10 influenza season. Available at: http://cdc.gov/h1n1flu/guidance/diagnostic_tests.htm. Accessed November 2009.
30. Centers for Disease Control and Prevention. Updated interim recommendations for the use of antiviral medications in the treatment and prevention of influenza for the 2009–2010 season. Available at: http://cdc.gov/h1n1flu/recommendations.htm. Accessed November 2009.
31. Centers for Disease Control and Prevention. Oseltamivir-resistant novel influenza A (H1N1) virus infection in two immunosuppressed patients–Seattle, Washington, 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 893–896.
32. Centers for Disease Control and Prevention. Oseltamivir-resistant 2009 pandemic influenza A (H1N1) virus infection in two summer campers receiving prophylaxis–North Carolina, 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 969–972.
33. Chen H, Cheung CL, Tai H, et al. Oseltamivir-resistant influenza A pandemic (H1N1) 2009 virus, Hong Kong, China. *Emerg Infect Dis*. doi://10.3201/eid1512.091057. Available at: http://www.cdc.gov.
34. World Health Organization. Pandemic (H1N1) 2009–update 70. Available at: http://www.who.int/csr/don/2009_10_16/en/index.html. Accessed November 2009.
35. Ciblak MA, Albayrak N, Odabas Y, et al. Cases of influenza A (H1N1) reported in Turkey, May-July 2009. *Euro Surveill* 2009; 14: 19304.
36. Centers for Disease Control and Prevention. Hospitalized patients with novel influenza A (H1N1) virus infection–California, April-May 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 536–541.
37. Centers for Disease Control and Prevention. Update: novel influenza A (H1N1) virus infection–Mexico, March-May, 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 585–589.
38. Centers for Disease Control and Prevention. Intensive-care patients with severe novel influenza A (H1N1) virus infection–Michigan, June 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 749–752.
39. Centers for Disease Control and Prevention. 2009 pandemic influenza A (H1N1) virus infection–Chicago, Illinois, April-July 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 913–918.
40. Gilsdorf A, Poggensee G. Influenza A(H1N1)v in Germany: the first 10,000 cases. *Euro Surveill* 2009; 14: 19318.
41. Health Protection Agency, Health Protection Scotland, National Public Health Service for Wales, and HPA Northern Ireland Swine Influenza Investigation Teams. Epidemiology of new influenza A (H1N1) virus infection, United Kingdom, April-June 2009. *Euro Surveill* 2009; 14: 19230.
42. New York City Department of Health and Mental Hygiene. 2009 New York City Department of Health and Mental Hygiene Health alert #27: pandemic (H1N1) 2009 influenza update, revised reporting requirements and testing procedures. Available at: http://www.nyc.gov/html/doh/downloads/pdf/cd/2009/09md27.pdf. Accessed November 2009.
43. Munayco CV, Gómez J, Laguna-Torres VA, et al. Epidemiological and transmissibility analysis of influenza A (H1N1)v in a Southern Hemisphere setting: Peru. *Euro Surveill* 2009; 14: 19299.
44. Perez-Padilla R, de la Rosa-Zamboni D, Ponce de Leon S, et al. Pneumonia and respiratory failure from swine-origin influenza A (H1N1) in Mexico. *N Engl J Med* 2009; 361: 680–689.
45. Vaillant L, Ruche G, Tarantola A, Barboza P. Epidemiology of fatal cases associated with pandemic H1N1 influenza 2009. *Euro Surveill* 2009; 14: 19309.
46. World Health Organization. Pandemic (H1N1) 2009–update 58. Available at: http://www.who.int/csr/don/2009_07_06/en/index.html. Accessed November 2009.
47. Centers for Disease Control and Prevention. Update on influenza A (H1N1) 2009 monovalent vaccines. *MMWR Morb Mortal Wkly Rep* 2009; 58: 1100–1101.
48. Centers for Disease Control and Prevention. Use of influenza A (H1N1) 2009 monovalent vaccine: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2009. *MMWR Morb Mortal Wkly Rep* 2009; 58: 1–8.
49. Zhu FC, Wang H, Fang HH, et al. A novel influenza A (H1N1) vaccine in various age groups. *N Engl J Med*. doi://10.1056/NEJMoa0908535. Available at: http://content.nejm.org.
50. National Institutes of Health. Early results: in children, 2009 H1N1 influenza vaccine works like seasonal flu vaccine. Available at: http://www.nih.gov/news/health/sep2009/niaid-21.htm. Accessed November 2009.
51. Ghendon YZ, Kaira AN, Elshina GA. The effect of mass influenza immunization in children on the morbidity of the unvaccinated elderly. *Epidemiol Infect* 2006; 134: 71–78.

Mount Sinai Journal of Medicine 77:114–123, 2010

SPECIAL FEATURE

Translating Knowledge About Environmental Health to Practitioners: Are We Doing Enough?

Leonardo Trasande, MD, MPP,[1,2] Nicholas Newman, DO,[3] Linda Long, MPH,[4] Genevieve Howe, MPH,[5] Beth J. Kerwin,[5] Richard J. Martin, DO, MPH,[6] Sheila A. Gahagan, MD, MPH,[3,7] and William B. Weil, MD[8]

[1]Department of Preventive Medicine, Mount Sinai School of Medicine, New York, NY
[2]Department of Pediatrics, Mount Sinai School of Medicine, New York, NY
[3]Pediatric Environmental Health and Lead Clinic, Cincinnati Children's Hospital Medical Center, Cincinnati, OH
[4]Michigan Chapter, American Academy of Pediatrics, Ann Arbor, MI
[5]Ecology Center and Michigan Network for Children's Environmental Health, Ann Arbor, MI
[6]La Clinica del Cariño Family Health Center, Hood River, OR
[7]Department of Pediatrics, University of California San Diego, San Diego, CA
[8]Department of Pediatrics, Michigan State University, East Lansing, MI

OUTLINE

METHODS
RESULTS
  Systematic Review
  Michigan Pediatrician Survey
CONCLUSIONS

ABSTRACT

Background: Practitioners see a large number of children affected by environmental exposures each year. A national network of pediatric environmental health specialty units has been established to strengthen prevention capacity, yet the effectiveness of that translational resource has not been assessed.

Methods: We supplemented a qualitative systematic review of previous assessments of healthcare provider capacity with a self-administered survey sent to the membership of the Michigan chapter of the American Academy of Pediatrics. We mailed surveys twice between October 2007 and January 2008 and obtained a 39.4% response rate.

Results: Our systematic review identified 8 relevant studies, all of which relied on self-report questionnaires and surveys. Recognizing this methodological weakness, we found that national and state samples consistently identified significant gaps in self-efficacy and knowledge about environmental hazards across a broad range of child care providers. In the Michigan survey, respondents voiced high self-efficacy in dealing with lead and second-hand smoke, but confidence in managing pesticide, mercury, mold, polychlorinated biphenyl, and air pollution exposures was much lower ($P < 0.0001$). Pediatricians routinely referred affected patients to lead/toxicology clinics and allergist/immunologists but not to the regional pediatric environmental health specialty unit.

Conclusions: Gaps persist in practitioner knowledge about environmental health nationwide and across disciplines. Despite methodological weaknesses, educational opportunities and other efforts should be

Address Correspondence to:

Leonardo Trasande
Children's Environmental
Health Center
Mount Sinai School of Medicine
New York, NY 10029
Email: leo.trasande@mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20158

studied to determine best practices for enhancing the evaluation of environmental health concerns in children. ***Mt Sinai J Med 77:114–123, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: asthma, children's environmental health, developmental disabilities, lead poisoning, practitioner knowledge, self-efficacy.

Children are uniquely vulnerable to many environmental chemicals that are widely distributed in the environment.[1–12] Aspects of children's behavior and biology that augment their vulnerability include high metabolic rates, developing organ systems, playing close to the ground, and hand-to-mouth behaviors.[13–15] The prevalence and incidence of asthma,[16] developmental disabilities,[17–19] certain childhood cancers,[20–23] and birth defects[24,25] have increased over the past 30 years across the United States. There is increasing evidence suggesting that chemical exposures contribute at least in part to these conditions.[23,26–32]

In response to increasing chronic disease of environmental origin in children and increasing understanding of opportunities for prevention, a network of translational centers called pediatric environmental health specialty units (PEHSUs) was established in 1998 with support from the Centers for Disease Control and Prevention and the Environmental Protection Agency (EPA). One such center was established in each of the 10 EPA regions nationwide to diagnose and treat children with diseases of environmental origin, to reduce environmental health threats to children, to improve practitioner access to expertise in environmental medicine, and to strengthen prevention capacity. These translational centers received a 2007 Children's Environmental Health Excellence Award from the EPA, and they have trained over 100,000 health professionals on children's environmental health issues.

> *In response to increasing chronic disease of environmental origin in children and increasing understanding of opportunities for prevention, a network of translational centers called pediatric environmental health specialty units (PEHSUs) was established in 1998 with support from the Centers for Disease Control and Prevention and the Environmental Protection Agency.*

Federal resources for public health programs are inherently scarce. Although the PEHSUs have been lauded by the EPA, it is important to document whether this translational resource has resolved gaps in clinician training or whether ongoing support is needed. We supplemented a qualitative systematic review of previous assessments of healthcare provider capacity with results from a survey that we collected in October 2007 from the membership of the Michigan chapter of the American Academy of Pediatrics (MIAAP).

METHODS

We executed the systematic review with PubMed as the data source and with *environmental health, children*, and *practice* as the search terms on May 11, 2009. We reviewed 820 abstracts to assess them for relevant content. In addition, 1 newly published letter to the editor[33] and 2 recently published articles[34,35] were also identified by the first author as eligible for inclusion in this analysis on the basis of their content, even though they were not yet listed on PubMed. In total, 9 publications were identified as assessing practitioner capacity with respect to children's environmental health.[33,36–43] Stevens *et al.*'s analysis[41] was excluded because it was a study of providers from the United Kingdom, as was Carlsson *et al.*'s article[35] because of its focus on Swedish providers.

To supplement the systematic review, we mailed a self-administered survey to the membership of the MIAAP twice between October 2007 and January 2008. Our survey was modeled on the one described by Kilpatrick *et al.*[40] and contained 22 questions divided into 3 sections. The first asked pediatricians to rate their agreement with a series of belief statements on a Likert scale of 1 to 5 (from "strongly disagree" to "strongly agree"). These questions asked pediatricians to evaluate their perceptions about the role of the environment in children's health, the need for environmental history taking, and their ability to control environmental exposures. For the purpose of this survey, we did not define *environment* for participants, and we defined *environmental health* as those aspects of human health and disease that are determined by factors in the environment and the theory and practice of assessing and controlling factors in the environment that can potentially affect health. Pediatricians were also asked to opine whether environmentally mediated disease in children was increasing and whether environmental history taking as a part of routine well-child care would take up too much time. The first section also ascertained respondent self-confidence in managing lead, pesticide, air pollution, second-hand smoke, mercury, mold, and polychlorinated biphenyl (PCB) exposures.

The second section of the survey asked pediatricians to quantify the number of patients in the past year (0, 1–5, 6–10, 11–15, or >15) affected by 9 categories of exposures (lead, mold, second-hand smoke, mercury, pesticide, poor air quality, poor water quality, carbon monoxide, and PCBs). A subsequent question identified elements of clinical practice regarding the treatment and evaluation of environmental health concerns (referral to lead/toxicology clinics, referral to an allergist/immunologist, self-management, or referral to other specialist). In the interest of brevity, this question was asked for environmental exposures in general and was not asked for each category of exposure.

We also inquired whether respondents had heard about the regional PEHSU and how many patients they had referred there in the past year. Respondents were asked to quantify the number of referrals that they would make annually if a pediatric environmental health clinic existed in Michigan (0, 1, 2–5, 6–10, 11–20, or >20). They were asked whether they owned a copy of the American Academy of Pediatrics handbook *Pediatric Environmental Health* and, if so, how often they referred to their book in clinical practice. They were asked whether they had received specific training in environmental history taking and whether they would be interested in additional training. The final section of the survey asked respondents for demographic and practice information.

We obtained approval for the survey from the MIAAP and from the institutional review board of the Michigan State University School of Medicine. A waiver of signed consent was granted. The institutional review board of the Mount Sinai School of Medicine granted L.T. an exemption from review for the purpose of analyzing the de-identified data. MIAAP mailed the questionnaire to all members not currently in training with a cover letter and a stamped, addressed return envelope in October 2007 and sent a second copy of the survey to those pediatricians who had not returned the surveys 3 months later. A follow-up e-mail reminder and an announcement in the MIAAP newsletter were also sent to the membership to encourage participation.

During data entry, we identified missing values and excluded those surveys from the analysis of the variable(s) for which the response was missing. We also checked data by running frequencies to check for outliers and data entry errors, and we randomly sampled and checked our data entry for 10% of the questionnaires to ensure accuracy. Descriptive data analyses were conducted with SAS 9.1.3 (Research Triangle Institute, Cary, NC). To assess patient volume for environmental health concerns, past referrals to the PEHSU, and potential referrals to a clinic for environmental concerns, a range was developed by multiplication of the lower and higher ends of the range by the number of respondents who selected an option (eg, 6 respondents reporting 2–5 referrals could have made 12–30 referrals). The lower and higher ends of patient volume reported by each of the respondents was summed to produce lower and higher end estimates for the actual volumes of affected patients, PEHSU referrals, and potential referrals to a pediatric environmental health clinic in Michigan.

RESULTS

Systematic Review

The 8 studies examined in this analysis are presented in Table 1. All of these studies relied on self-reporting. Methodological approaches such as chart reviews, responses to programmed patients, or responses to case descriptions were not identified through the systematic review. Although self-report questionnaires and surveys were the only methodologies identified, national and state samples consistently identified significant gaps in self-efficacy and knowledge about environmental hazards across a broad range of child care providers. Surveys in 3 states suggested demand for additional clinical resources in children's environmental health to fill these gaps.

> *Although self-report questionnaires and surveys were the only methodologies identified, national and state samples consistently identified significant gaps in self-efficacy and knowledge about environmental hazards across a broad range of child care providers.*

Table 1. *Publications Identified Through a Systematic Review.*

Reference	Study Design	Respondent Sample Characteristics and Size	Response Rate	Findings
30	Survey, self-report	1396 New York, Connecticut, and New Jersey pediatricians who were American Academy of Pediatrics members	36% combined weighted rate (26% by mail, 43% by e-mail)	21% of respondents found practices affected by the attacks of September 11 and were often asked about air quality and environmental toxins; 86% reported that training left them ill prepared to handle these issues.
31	Survey, self-report	277 New York American Academy of Pediatrics members	20.3%	Respondents agreed that the role of the environment in children's health is significant (mean, 4.14 on 1–5 Likert scale). They had high self-efficacy in dealing with lead exposure (mean, 4.16–4.24), but their confidence in their skills for addressing pesticides, mercury, and mold was much lower (mean, 2.51–3.21; $P<0.001$).
32	Survey, self-report	277 Wisconsin American Academy of Pediatrics members	35.4%	Results were similar to those of reference 31. They rarely referred to the pediatric environmental health specialty unit for the area; 88.6% would refer patients to a clinic "where pediatricians could refer patients for clinical evaluation and treatment of their environmental health concerns."
33	Survey, self-report	177 members of the Minnesota American Academy of Pediatrics	24.4%	Results were similar to those of reference 31. They rarely referred to the pediatric environmental health specialty unit for the area; 87.7% would refer patients to an environmental specialty clinic.
34	Survey, self-report	266 Georgia American Academy of Pediatrics members	55.8%	53.5% of respondents reported experience with a patient who was seriously affected by an environmental exposure. Fewer than 1 in 5 reported having received training in environmental history taking.
36	Survey, self-report	1735 family practitioners, pediatricians, naturopaths, school nurses, public health officials, Head Start parents, and parent-teacher association parents from the Northwest United States	24%	There were consistent concerns about environmental exposures in general; there was variability in the degree of concern about specific exposures, especially among clinicians, parents, and public health officers.

(*continued overleaf*)

Table 1. *(Continued).*

Reference	Study Design	Respondent Sample Characteristics and Size	Response Rate	Findings
37	Survey, self-report	882 US American Academy of Pediatrics members	54.6%	Many pediatricians believe sun exposure is worthy of concern, a minority provide counseling, and 38% rated sun exposure counseling as very important in comparison with car seats (86%) and nutrition (74%)
38	Survey, self-report	25 attendees of the meeting of the Environmental Health Special Interest Group of the Pediatric Academic Societies	50%	All faculty respondents stated that pediatric resident training in environmental health is insufficient. There was strong confidence in lead, asthma, tobacco smoke, and skin cancer prevention, but there was much less confidence about pesticide exposure, volatile organic compounds, and food contaminants.

Michigan Pediatrician Survey

Of the 989 questionnaires mailed, 383 were returned, and 17 were undeliverable (overall response rate = 383/972 or 39.4%). After the exclusion of 21 of the 383 returned questionnaires (because of incomplete data or an indication by the respondent that he or she was no longer practicing), the final analyzed sample consisted of 362 practicing pediatricians. Because surveys were distributed directly by the MIAAP, nonresponse analysis was limited to a comparison of respondents and nonrespondents by the first 3 digits of their zip codes.

A description of respondents is provided in Table 2. The respondents were 59.5% women, and nearly four-fifths were in primary care. Private and public/community clinic settings predominated, and various specialties were also represented, including neonatology, developmental/behavioral pediatrics, and hematology/oncology. About one-fifth of the respondents reported a majority of patients on public assistance. Every region in the state corresponding to the first 3 digits of the respondent's zip code was represented in the survey. Respondents and nonrespondents did not significantly differ by the first 3 digits of the zip code ($P = 0.999$).

Pediatricians agreed strongly that the role of the environment in health is great (Table 3), that assessing environmental exposure through history taking is important, and that the magnitude of environmentally mediated illness is increasing. Respondents voiced high levels of self-confidence

Table 2. *Description of Michigan Respondents and Their Practices.*

Characteristic	n	%
Age (mean ± SD)	49.3 ± 12.1	
Years in practice (mean ± SD)	17.8 ± 11.8	
Sex		
Male	142	40.2
Female	211	59.8
Practice type		
Primary care	265	76.6
Specialty	69	19.9
Urgent care/emergency	12	3.5
Practice setting		
Private, not exclusively with HMO	186	52.8
Public/community clinic	115	32.7
Teaching	24	6.8
Other	15	4.3
Private, exclusively with HMO	9	2.6
Research	3	0.9
Patients on Medicaid or public assistance		
0%–24%	152	44.8
25%–49%	127	25.7
50%–74%	67	19.8
75%–100%	33	9.7
Urban area (by a 3-digit zip code)		
Urban	176	54.0
Nonurban	150	46.0

Abbreviations: HMO, health maintenance organization; SD, standard deviation.

Pediatricians agreed strongly that the role of the environment in health is great, that assessing environmental exposure through history taking is important, and that the magnitude of environmentally mediated illness is increasing.

for lead (mean, 4.06) and second-hand smoke exposure (mean, 4.07). However, for pesticide, air pollution, PCB, mercury, and mold exposures, the mean responses were much lower (all $P < 0.001$) in comparison with lead and second-hand smoke, ranging from 1.86 to 2.71.

Almost all the respondents (95.4%) reported a past experience with a patient affected by an environmental exposure; 81.7% reported past experience with patients affected by exposures other than second-hand smoke. Second-hand smoke was the most common clinical experience (Table 4); it was followed by mold, lead, and poor air quality. Although the majority of the patient volume was for exposures for which respondents described strong self-confidence, respondents reported 1054 to 1470 patients affected by poor air quality and 1171 to 1955 patients affected by mold, and low self-confidence was reported for exposures to these 2 factors. Patients affected by poor water quality and pesticides numbered in the hundreds. Relatively few pediatricians (Table 5; 6.0%) knew about or made referrals (0.9%) to the PEHSU, and no respondent made more than 1 referral. Typical management approaches included referral to a local toxicology/lead clinic (41.9%), management without further assistance (35.9%), and referral to an allergist/immunologist (26.0%). No respondents included the PEHSU as part of routine management.

Relatively few pediatricians (6.0%) knew about or made referrals (0.9%) to the PEHSU, and no respondent made more than 1 referral.

Fewer than one-sixth of respondents reported any training in environmental history taking (15.4%), and many (78.1%) voiced interest in additional training. Grand rounds presentations and web-based seminars were the most strongly preferred approaches for this training. Fewer than half of the respondents (39.0%) had a copy of the American Academy of Pediatrics text on pediatric environmental health, and of those who owned a copy, the vast majority used it yearly or once every few months.

Despite the relatively low referral rate to the PEHSU, 89.6% of respondents would refer patients to a clinic "where pediatricians could refer patients for clinical evaluation and treatment of their environmental health concerns." The 276

Table 3. *Pediatricians' Self-Reported Beliefs and Self-Efficacy Regarding Environmental Health in Four States.*

Belief Statements	Michigan Mean ± SD
The role of environmental health impacts on children is (1) of little importance → (5) of great importance (n = 361).	4.42 ± .71
The amount of control pediatricians have over environmental health hazards is (1) minimal → (5) maximal (n = 358).	2.62 ± .94
The magnitude of children's environment-related illnesses is (1) decreasing → (5) increasing (n = 357).	3.76 ± .88
Assessing environmental exposures through history taking in pediatric practice is (1) of little importance → (5) of great importance (n = 358).	3.96 ± .88
Conducting an environmental health history on all my patients (1) takes up too much time → (5) does not take up too much time (n = 354).	3.08 ± .99
Self-Efficacy Statements: Confidence in Management	Michigan Mean ± SD
Lead exposure (n = 357)	4.06 ± 1.03
Pesticide exposure (n = 355)	2.48 ± 1.11*
Air pollution exposure (n = 354)	2.60 ± 1.10*
Second-hand smoke (n = 357)	4.06 ± 0.99
Mercury exposure (n = 355)	2.49 ± 0.99*
Mold exposure (n = 356)	2.71 ± 1.13*
PCB exposure (n = 356)	1.87 ± 0.89*

Abbreviations: PBC, polychlorinated biphenyl; SD, standard deviation.
*$P < 0.0001$ in comparison with lead and second-hand smoke.

Table 4. *Frequencies of Pediatrician Activities Regarding Environmental Health.*

Clinical Activity	n (%)
Knows about the PEHSU (n = 333)	21 (6.1)
Refers to the PEHSU (n = 331)	3 (0.9)
Has received environmental history training (n = 342)	52 (15.2)
Is interested in more training in environmental health (n = 340)	248 (73.0)
Current practice regarding the clinical evaluation and treatment of pediatric environmental health concerns:	
Referral to PEHSU (n = 351)	0 (0.0)
Referral to a local toxicology/lead clinic (n = 351)	153 (42.4)
Referral to an allergist/immunologist (n = 351)	96 (26.6)
Referral to other specialist (n = 351)	63 (17.5)
Management of the patient by the pediatrician himself (n = 351)	128 (35.5)
Would refer patients to a referral clinic for the evaluation and treatment of pediatric environmental health concerns (n = 317):	283 (89.3)
Would refer 1 patient per year	29 (9.4)
Would refer 2–5 patients per year	146 (46.1)
Would refer 6–10 patients per year	63 (19.9)
Would refer 11–20 patients per year	29 (9.2)
Would refer more than 20 patients per year	16 (5.1)
Expected total volume of referrals to a clinic for pediatric environmental health concerns	1338–2289

Abbreviation: PEHSU, pediatric environmental health specialty unit.

Table 5. *Frequencies of Respondents Noting Clinically Significant Concerns.*

Clinical Concern	No Patients	1–10 Patients Affected in the Past Year	11 or More Patients Affected in the Past Year	Estimate of Total Patients Affected in the Past Year
Lead (n = 346)	143 (41.3%)	174 (50.3%)	29 (8.4%)	710–1430
Mold (n = 337)	102 (30.3%)	181 (53.7%)	54 (16.0%)	1171–1955
Carbon monoxide (n = 335)	226 (67.4%)	108 (32.2%)	0 (0%)	153–585
Second-hand smoke (n = 343)	23 (6.7%)	56 (16.3%)	264 (77.0%)	4079–4415
Mercury (n = 335)	305 (91.0%)	28 (8.4%)	2 (0.6%)	58–170
Poor water quality (n = 333)	246 (73.9%)	75 (22.5%)	12 (3.6%)	298–610
Poor air quality (n = 329)	182 (55.3%)	85 (25.8%)	62 (18.8%)	1054–1470
Pesticides (n = 333)	282 (84.7%)	48 (14.4%)	3 (0.9%)	124–320
PCBs or other organic compounds (n = 330)	302 (91.5%)	26 (7.9%)	2 (0.6%)	62–170

Abbreviation: PCB, polychlorinated biphenyl.

respondents would make as many as 1338 to 2239 referrals to such a clinical resource.

CONCLUSIONS

Michigan pediatricians, like their counterparts in other national studies and in studies in individual states in different regions of the country,[36–40,42,43] demonstrated strong interest in children's environmental health, agreed that environmental exposures affect their patients' health, and were interested in additional training. They also perceived environmental health concerns to be significant and common in their practices. Except for lead and second-hand smoke exposure, pediatricians reported little prior training in taking environmental histories and low self-confidence in managing these exposures.

Our response rate of 39.4% is low, and our results may reflect significant selection bias. However, this rate is typical for a physician mail survey,[37,39,44,45] and by the first 3 digits of their zip codes, nonrespondents and respondents were roughly similar. Nonrespondents may have lower levels of concern about environmental health problems than respondents, and respondents may be motivated to give the "correct" answer and overstate their level of interest. Nevertheless, we believe that our results indicate considerable interest in pediatric environmental health and opportunities for educating pediatricians about prevention.

The strong demand for a pediatric environmental health referral resource persisted despite self-reported use of existing toxicology/lead programs, and this suggests that a gap remains. For lead and second-hand smoke, a high volume of affected patients was reported, but respondents reported high efficacy, suggesting that they could manage these exposures well. Yet respondents also reported patients affected by exposures for which respondents had low self-confidence. Pediatricians did frequently refer to toxicologists, lead experts, and allergists/immunologists, and these specialists can and likely do manage a significant number of affected patients in Michigan very well. Toxicologists are trained to focus on the acute effects of mercury, pesticides, and other chemicals, but they are not specifically trained in the prevention of indolent, low-dose exposures to these toxicants and others, such as pesticides and poor air quality. Although allergists/immunologists can guide the treatment of mold-associated conditions such as asthma, remediation typically requires correction of the underlying humidity that fostered the development of the mold in the first place. The involvement of industrial hygienists and occupational/environmental medicine specialists can be extremely useful in the safe and effective removal of mold from affected areas in the home.

> *The strong demand for a pediatric environmental health referral resource persisted despite self-reported use of existing toxicology/lead programs, and this suggests that a gap remains.*

The regional PEHSU, located in Chicago, unites the expertise of these specialists, but we noted a near absence of referrals from Michigan pediatricians. Lack of familiarity appears to be a major factor, but the relative distance from Chicago may transform the PEHSU into a telephone resource, which may be less valuable. Similarly, in the New York State pediatrician survey, all of the reported referrals came from the southern part of the state that is proximate to the PEHSU in New York City.[37] Wisconsin and Minnesota are also served by the Chicago PEHSU, and distance may have explained the infrequent referral rate to the PEHSU from those states.[38,39] Pediatricians may prefer a nearby clinical site at which an in-person evaluation would occur, just as they might prefer to refer patients to a local cardiologist rather than one at a more distant institution.

Our analysis does not definitively suggest that Michigan should establish additional clinical resources for affected patients. The need for the additional use of referral resources is well established, but the impact of marketing or public relations to increase the utilization of the existing environmental health specialty units needs to be studied before the establishment of more centers. Additional grand rounds presentations and continuing medical education sessions in Michigan may go far toward improving pediatrician knowledge about environmental hazards and their management in children. A poison control network already exists in Michigan, and clinics established to manage lead hazards could be enhanced to respond to a broader array of concerns. Pediatric environmental health has potential scenarios in which clinical evaluations may be well served by telephone and electronic communication, and in-person access may not be seriously needed. This begs the question whether existing centers could not meet much of the need described in this article.

The results of this survey and the systematic review further underline a number of gaps in the protection of children from environmental hazards. Data from the Michigan survey and the Minnesota, New York, and Wisconsin surveys that we reviewed suggest infrequent use of the PEHSUs despite high need. Together, these studies suggest that additional efforts are needed to increase familiarity with the PEHSUs. Survey data from New York have spurred efforts to develop a regionalized system of centers of excellence in environmental pediatrics there (A. Carlson, New York State Department of Health, personal communication, 2008), and future efforts should focus on the impact of different approaches pursued across the country to compensate for the minimal attention that environmental health receives in the medical school curriculum.[46]

ACKNOWLEDGMENT

This research was supported by the Beldon Fund. The authors thank the many pediatricians who took the time to respond to the survey.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Laraque D, Trasande L. Lead poisoning: successes and 21st century challenges. *Pediatr Rev* 2005; 26: 435–443.
2. Canfield RL, Henderson Jr CR, Cory-Slechta DA, et al. Intellectual impairment in children with blood lead concentrations below 10 µg per deciliter. *N Engl J Med* 2003; 348: 1517.
3. Trasande L, Landrigan PJ, Schechter C. Public health and economic consequences of methyl mercury toxicity to the developing brain. *Environ Health Perspect* 2005; 113: 590–596.
4. Grandjean P, Budtz-Jorgensen E, White RF, et al. Methylmercury exposure biomarkers as indicators of neurotoxicity in children aged 7 years. *Am J Epidemiol* 1999; 150: 301–315.
5. Grandjean P, Weihe P, White RF, et al. Cognitive deficit in 7-year-old children with prenatal exposure to methylmercury. *Neurotoxicol Teratol* 1997; 19: 417–428.
6. Whyatt RM, Rauh V, Barr DB, et al. Prenatal insecticide exposures and birth weight and length among an urban minority cohort. *Environ Health Perspect* 2004; 112: 1125–1132.
7. Weitzman M, Cook S, Auinger P, et al. Tobacco smoke exposure is associated with the metabolic syndrome in adolescents. *Circulation* 2005; 112: 862–869.
8. Weitzman M, Gortmaker S, Walker DK, Sobol A. Maternal smoking and childhood asthma. *Pediatrics* 1990; 85: 505–511.
9. Lupton C, Burd L, Harwood R. Cost of fetal alcohol spectrum disorders. *Am J Med Genet* 2004; 127: 42–50.
10. Jacobson JL, Jacobson SW. Intellectual impairment in children exposed to polychlorinated biphenyls in utero. *N Engl J Med* 1996; 335: 783.
11. Trasande L, Thurston GD. The role of air pollution in asthma and other pediatric morbidities. *J Allergy Clin Immunol* 2005; 115: 689–699.
12. Romieu I, Meneses F, Ruiz S, et al. Effects of air pollution on the respiratory health of asthmatic children living in Mexico City. *Am J Respir Crit Care Med* 1996; 154: 300–307.
13. National Research Council. *Pesticides in the Diets of Infants and Children*. Washington, DC: National Academy Press; 1993.
14. Freeman NCG, Hore P, Black K, et al. Contributions of children's activities to pesticide hand loadings following residential pesticide application. *J Expo Sci Environ Epidemiol* 2004; 15: 81–88.
15. Gurunathan S, Robson M, Freeman N, et al. Accumulation of chlorpyrifos on residential surfaces and toys accessible to children. *Environ Health Perspect* 1998; 106: 9–16.
16. Mannino DM, Homa DM, Akinbami LJ, et al. Surveillance for asthma–United States, 1980–1999. *MMWR Surveill Summ* 2002; 51: 1–13.
17. Newschaffer CJ, Falb MD, Gurney JG. National autism prevalence trends from United States special education data. *Pediatrics* 2005; 115: e277–e282.
18. Zito JM, Safer DJ, dosReis S, et al. Trends in the prescribing of psychotropic medications to preschoolers. *JAMA* 2000; 283: 1025–1030.
19. LeFever GB. The extent of drug therapy for attention deficit-hyperactivity disorder among children in public schools. *Am J Public Health* 1999; 89: 1359–1364.
20. Schechter CB. Re: Brain and other central nervous system cancers: recent trends in incidence and mortality. *J Natl Cancer Inst* 1999; 91: 2050–2051.
21. Devesa SS, Blot WJ, Stone BJ, et al. Recent cancer trends in the United States. *J Natl Cancer Inst* 1995; 87: 175–182.
22. Infante-Rivard C, Sinnett D. Preconceptional paternal exposure to pesticides and increased risk of childhood leukaemia. *Lancet* 1999; 354: 1819.
23. Infante-Rivard C, Weichenthal S. Pesticides and childhood cancer: an update of Zahm and Ward's 1998 review. *J Toxicol Environ Health B Crit Rev* 2007; 10: 81–99.
24. Paulozzi LJ, Erickson JD, Jackson RJ. Hypospadias trends in two US surveillance systems. *Pediatrics* 1997; 100: 831–834.
25. Williams LJ, Kucik JE, Alverson CJ, et al. Epidemiology of gastroschisis in metropolitan Atlanta, 1968 through 2000. *Birth Defects Res A Clin Mol Teratol* 2005; 73: 177–183.
26. Kattan M, Stearns SC, Crain EF, et al. Cost-effectiveness of a home-based environmental intervention for inner-city children with asthma. *J Allergy Clin Immunol* 2005; 116: 1058–1063.
27. McConnell R, Berhane K, Gilliland F, et al. Asthma in exercising children exposed to ozone: a cohort study. *Lancet* 2002; 359: 386–391.
28. Friedman MS, Powell KE, Hutwagner L, et al. Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma. *JAMA* 2001; 285: 897–905.
29. National Academy of Sciences Committee on Developmental Toxicology. *Scientific Frontiers in Developmental Toxicology and Risk Assessment*. Washington, DC: National Academies Press; 2000.
30. Knox EG. Childhood cancers and atmospheric carcinogens. *Br Med J* 2005; 59: 101.
31. Lee WJ, Cantor KP, Berzofsky JA, et al. Non-Hodgkin's lymphoma among asthmatics exposed to pesticides. *Int J Cancer* 2004; 111: 298–302.
32. Daniels JL, Olshan AF, Teschke K, et al. Residential pesticide exposure and neuroblastoma. *Epidemiology* 2001; 12: 20–27.
33. Roberts JR, Balk SJ, Forman J, Shannon MW. Teaching about pediatric environmental health. *Ambul Pediatr* 2009; 9: 129–130.
34. Rogers B, McCurdy LE, Slavin K, et al. Children's Environmental Health Faculty Champions Initiative: a successful model for integrating environmental health into pediatric healthcare. *Environ Health Perspect* 2009; 117: 850–855.
35. Carlsson N, Johansson A, Hermansson G, Andersson-Gäre B. Child health nurses' roles and attitudes in reducing children's tobacco smoke exposure. *J Clin Nurs*. doi:/10.1111/j.1365–2702.2009.02847.x. Available at: http://www3.interscience.wiley.com.
36. Hu YY, Adams RE, Boscarino JA, Laraque D. Training needs of pediatricians facing the environmental health and bioterrorism consequences of September 11th. *Mt Sinai J Med* 2006; 73: 1156–1164.
37. Trasande L, Boscarino J, Graber N, et al. The environment in pediatric practice: a study of New York pediatricians' attitudes, beliefs, and practices towards children's environmental health. *J Urban Health* 2006; 83: 760–772.

38. Trasande L, Schapiro ML, Falk R, et al. Pediatrician attitudes, clinical activities, and knowledge of environmental health in Wisconsin. *WMJ* 2006; 105: 45–49.
39. Trasande L, Ziebold C, Schiff JS, et al. The role of the environment in pediatric practice in Minnesota: attitudes, beliefs, and practices towards children's environmental health. *Minn Med* 2008; 91: 36–39.
40. Kilpatrick N, Frumkin H, Trowbridge J, et al. The environmental history in pediatric practice: a study of pediatricians' attitudes, beliefs, and practices. *Environ Health Perspect* 2002; 110: 823–827.
41. Stevens E, Cullinan P, Colvile R. Urban air pollution and children's asthma: what do parents and health professionals think? *Pediatr Pulmonol* 2004; 37: 530–536.
42. Chai SK, Robertson WO, Takaro T, et al. Pediatric environmental health: perceptions from a survey questionnaire. *Int J Occup Environ Health* 2001; 7: 37–43.
43. Balk SJ, O'Connor KG, Saraiya M. Counseling parents and children on sun protection: a national survey of pediatricians. *Pediatrics* 2004; 114: 1056–1064.
44. Kellerman SE, Herold J. Physician response to surveys: a review of the literature. *Am J Prev Med* 2001; 20: 61–67.
45. Trasande L, Cronk CE, Leuthner SR, et al. The National Children's Study and the children of Wisconsin. *WMJ* 2006; 105: 50–54.
46. Szneke P, Nielsen C, Tolentino N. Connecticut physicians' knowledge and needs assessment of environmentally related health hazards–a survey. *Conn Med* 1994; 58: 131–135.

SPECIAL FEATURE

Surgical Treatment of Metastatic Spinal Tumors

Scott A. Meyer, MD, Harshpal Singh, MD, and Arthur L. Jenkins, MD, FACS

Department of Neurosurgery, Mount Sinai School of Medicine, New York, NY

OUTLINE

Surgical Treatment
Metastatic Epidural Spinal Cord Compression
Spinal Instability and Deformity
Minimally Invasive Approaches
Conclusions

ABSTRACT

Spinal metastases represent a significant cause of morbidity in patients diagnosed with malignancies. Metastases to the spine can cause severe pain, paralysis, and impairment of activities of daily living. The treatment paradigm for spinal metastases involves a cohesive multidisciplinary approach that allows treatment plans to be made in the context of a patient's overall condition. There have been significant advances in the surgical treatment of spinal metastases that can significantly improve a patient's quality of life. ***Mt Sinai J Med 77:124–129, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: epidural metastasis, pathological fracture, spinal cord compression, spinal instability, spinal metastasis.

Metastases to the spine are commonly seen in patients who harbor malignancies, and 10% of all spinal metastases have no known primary tumor at the time of presentation.[1] Autopsy series have demonstrated that vertebral metastases are found in as many as 30% of patients with malignancies.[2] Spinal metastases are most commonly seen, in decreasing order of incidence, from primary lung, prostate, kidney, liver, gastric, and colon cancers in males and primary breast, lung, uterine, thyroid, and gastric cancers in females.[3] In addition, there is a significant contribution from hematogenous disorders such as lymphoma and multiple myeloma. Metastases to the spine can cause severe pain, paralysis, and impairment of activities of daily living.[4] Magnetic resonance imaging is the diagnostic modality of choice in the diagnosis of spinal metastases, with a sensitivity of 93% and a specificity of 97%.[5] Additional information regarding the extent of bone destruction and alignment can be obtained with computed tomography of the spine.

We have found that a cohesive multidisciplinary approach to the treatment of metastases to the spine can improve outcomes as it brings the treating physicians together and allows plans to be made in the context of the patient's overall condition. The initial evaluation should involve a thorough history, including any prior history of malignancy. In some cases, symptoms related to spinal metastases are the first signs of malignancy, and a thorough metastatic workup, including standard X-rays, computed tomography of the chest, abdomen, and pelvis, positron emission tomography, and bone scans, is often necessary to identify the primary tumor. In addition, blood dyscrasias such as multiple myeloma and lymphoma need to be considered, and this requires further laboratory studies of the blood and possibly cerebrospinal fluid. Treatment paradigms involve close collaboration between the medical oncologist, radiation oncologist, and spine surgeon. The pathology of the primary tumor is important in determining the patient's life expectancy, sensitivity to adjuvant treatments, and likelihood of achieving systemic control.

Address Correspondence to:

Arthur L. Jenkins
Department of Neurosurgery
Mount Sinai School of Medicine
New York, NY
Email: arthur.jenkins@mountsinai.org

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20162

SURGICAL TREATMENT

The surgical treatment of spinal metastases has evolved significantly over the last couple of decades, with mounting evidence of the value of surgery in the treatment armamentarium. Decades ago, the role of surgery was limited to posterior laminectomy for neurological decompression in the case of metastatic epidural spinal cord compression (MESCC). The results were mixed, and with the advent of radiotherapy, surgery was all but abandoned when several retrospective studies[6–9] and a small randomized trial[10] did not show any benefit from laminectomy alone or in combination with radiotherapy in comparison with radiotherapy alone. This is not surprising given the limitations of laminectomy in addressing an anterior pathology. Approximately 85% of spinal metastases are a result of hematogenous spread to the vertebral body[11] and are thus not directly addressed in posterior-only approaches. Laminectomy performed for a pathology that involves the vertebral body will not result in a safe, direct decompression of the tumor. In addition, the indirect posterior decompression from laminectomy can lead to a "bowstring" of the spinal cord, stretching it across the anterior pathology.

However, over the last couple of decades, significant advances in technique and spinal instrumentation have led to improved patient outcomes. The surgical treatment paradigm has changed significantly, with movement from laminectomy alone to anterior decompression with instrumented stabilization. Although laminectomy helps to increase the diameter of the spinal canal at the affected levels, it does not achieve immediate decompression of the spinal cord and can lead to instability without instrumentation. The instability produced by laminectomy alone is compounded by the frequent loss of anterior column support in patients with MESCC, which, combined with the loss of posterior tension band support, can lead to stretching of the spinal cord across the anterior pathology and poor clinical outcomes. The surgical approach largely depends on the anatomic position of the metastasis. Tumors anterior to the spinal cord in the cervical spine are usually approached through an anterior approach, with decompression and stabilization possible. Anterior approaches in the thoracic and lumbosacral spine are more challenging because they may necessitate a transcavitary or posterolateral approach. However, these approaches are often well tolerated. Increasingly, anterior decompression through a posterolateral or lateral approach is becoming popular. These approaches allow for anterior decompression with anterior column support as well as posterior instrumented fusion for improved stability. Currently, the role of surgery can be divided into cancer treatment through oncological resection, neurological palliation, stabilization, pathological diagnosis, and pain relief.[2] In select cases, oncological resection of a spinal metastasis can be performed. When an oncological resection is possible, it can provide the best opportunity for long-term local control and palliation in patients with isolated metastases or radioresistant tumors. The goal in these select circumstances should be to achieve negative margins.[2,12] In the case of a solitary vertebral metastasis, this can be achieved through total or subtotal spondylectomy.[13]

> *Surgical options in the treatment of metastatic spinal tumors include] oncological resection, neurological palliation, stabilization, pathological diagnosis, and pain relief.*

Palliation is also an important role for the surgical management of spinal metastases. Improvements in approaches and anesthesia management have resulted in reduced morbidity. Advances in segmental spinal fixation systems have greatly enhanced the surgeon's ability to stabilize the spine sometimes with shorter segmental fixation. Various cage systems exist that allow the spinal surgeon to provide anterior column support after vertebrectomy. Poly(methyl methacrylate) (PMMA) has long been a useful adjunct treatment in providing immediate anterior column support in patients with lytic processes or after resections including the vertebral body.

METASTATIC EPIDURAL SPINAL CORD COMPRESSION

Metastases to the vertebrae can produce compression of the spinal cord leading to acute deterioration that requires urgent attention and treatment. MESCC occurs in approximately 5% of patients with cancer[14] (Figure 1). Patients who present with paraparesis from MESCC will likely become paraplegic if untreated. Most cases of MESCC occur in patients with a previously diagnosed malignancy, although it can be the initial manifestation in approximately 20% of patients.[15,16] This is particularly the case with

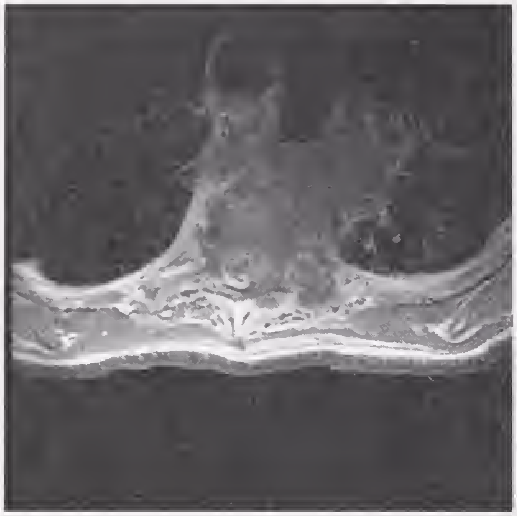


Fig 1. Axial magnetic resonance imaging demonstrating a metastatic epidural tumor involving the left pedicle and vertebral body and leading to significant spinal canal narrowing and spinal cord compression.

Most cases of metastatic epidural spinal cord compression occur in patients with a previously diagnosed malignancy, although it can be the initial manifestation in approximately 20% of patients.

lung cancer, for which MESCC is the presenting manifestation of malignancy in 30% of cases.[15,16] Prostate cancer, breast cancer, and lung cancer each account for approximately 15% to 20% of MESCC cases; non-Hodgkin's lymphoma, renal cell carcinoma, and multiple myeloma account for 5% to 10%, with the remainder coming from colorectal, sarcoma, or unknown primary tumors.[1,14,15,17] MESCC occurs in the thoracic spine in 60%, in the lumbosacral spine in 25%, and in the cervical spine in 15%, and this reflects the relative distribution of blood flow and bone mass in each of the spinal segments.[1,14,15]

Back pain is present in 95% of patients at the time of diagnosis and is typically but not universally worse at night or when the patient is recumbent.[15,17] Pain made worse with movement is suggestive of mechanical instability and is classically associated with pathological compression fractures and spinal deformity. Weakness is the second most common symptomatic manifestation of MESCC. There is a strong association between weakness and a patient's ambulatory status, with 50% to 68% of patients unable to walk at the time of diagnosis of MESCC.[1,6,14,15,17] The neurological examination should include an evaluation for sensory deficits, upper motor neuron signs, and autonomic dysfunction. Although isolated bowel or bladder symptoms are rarely the presenting symptoms,[1] 50% to 60% of patients have bowel or bladder symptoms at the time of diagnosis.[6,14]

The treatment armamentarium for MESCC includes corticosteroids, surgery, and radiotherapy along with aggressive systemic management. The optimal loading dose and maintenance dose of corticosteroids have not yet been determined. However, the beneficial effects of steroids on patients with MESCC have been demonstrated. Sorensen and colleagues[9] found that among patients undergoing radiotherapy for MESCC, those who were placed on a regimen of dexamethasone had statistically significantly improved rates of ambulation at 3 and 6 months. Radiotherapy has played an important role in the treatment of MESCC and will likely remain the primary treatment modality for patients with certain highly radiosensitive tumors. Research continues on the optimal dose, dose schedule, and field to achieve maximal tumor reduction while minimizing injury to sensitive adjacent structures such as the spinal cord. Although the surgical management of spinal metastases has improved, radiotherapy remains the primary treatment modality in many cases. Radiotherapy to the postoperative bed also plays an important role in the adjuvant therapy of patients who undergo surgical decompression.

A recent randomized trial by Patchell and colleagues[18] found that in patients who met certain criteria, surgical decompression followed by radiotherapy was superior to radiotherapy alone. The study included patients with a metastasis to a single area, which could include several contiguous vertebral segments, and importantly excluded patients with lymphoma, leukemia, multiple myeloma, and germ cell tumors. The study was stopped early during an interim analysis because of superior results in the surgery and radiotherapy group compared to the radiotherapy alone group. More patients regained the ability to walk after treatment and retained the ability to walk for a longer period of time in the surgical group than in the radiotherapy alone group.[18]

A recent randomized trial by Patchell and colleagues[18] found that in patients who met certain criteria, surgical decompression followed by radiotherapy was superior to radiotherapy alone.

In addition, there was improved pain control in the surgical group with decreased opioid analgesic requirements.[18] The difference in the improved outcome appears to be related to age, with the benefits of surgery with radiation over radiation alone disappearing in patients over the age of 60 years.[19]

SPINAL INSTABILITY AND DEFORMITY

Many metastatic lesions involving the spine are destructive in nature and can lead to bony destruction, ligamentous instability, and, in some cases, significant spinal deformity. The approach to patients with spinal deformities incorporates many of the basic tenets of spinal stability but has several unique characteristics. For one, the patient population as a whole tends to have more medical comorbidities; as such, approaches have to be well tolerated but also achieve the desired result. The preference of the metastasis for the vertebral body may lead to pathological compression fractures, and it often involves both the anterior and middle columns. The destruction of the vertebral body leads to sagittal plane abnormalities, with kyphosis being the predominant deformity in metastatic disease. The segmental instability produced by tumor-induced destruction of the vertebral body can lead to pathological biomechanical load sharing and result in significant pain.

Many metastatic lesions involving the spine are destructive in nature and can lead to bony destruction, ligamentous instability, and, in some cases, significant spinal deformity.

Advances in instrumentation technology have enabled spine surgeons to provide greater stability by using pedicle screw fixation, and oncological deformity can often be corrected as it tends to be flexible.[20] However, because a vertebrectomy is often performed as part of the decompression, reconstruction of the anterior and middle columns may be required to achieve stability (Figure 2). There are several options available for anterior reconstruction, including autografts, allografts, PMMA, and spacers.[20] It is important to remember that anterior reconstruction alone may not be adequate for support if the posterior ligamentous complex is involved in the case of patients with kyphosis.

Advances in instrumentation technology have enabled spine surgeons to provide greater stability by using pedicle screw fixation, and oncological deformity can often be corrected as it tends to be flexible.

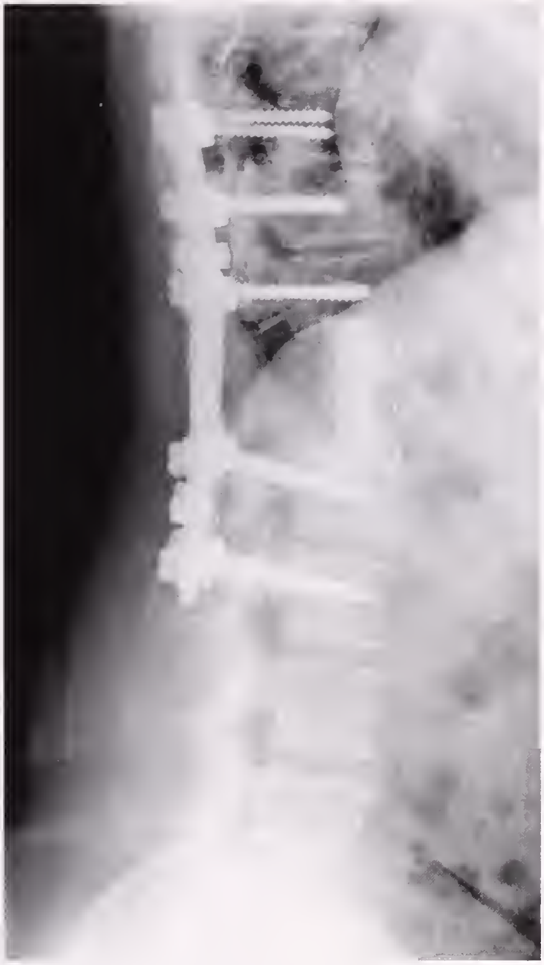


Fig 2. Plain lateral X-ray of the thoracolumbar spine demonstrating anterior column reconstruction with a titanium cage and posterior segmental pedicle screw fixation.

MINIMALLY INVASIVE APPROACHES

Recently, minimally invasive techniques have been increasingly used in patients because of the potential benefits of decreased pain, decreased blood loss, and smaller incisions. Minimally invasive transpedicular approaches to metastatic disease can palliate the symptoms of a patient with MESCC, leading to

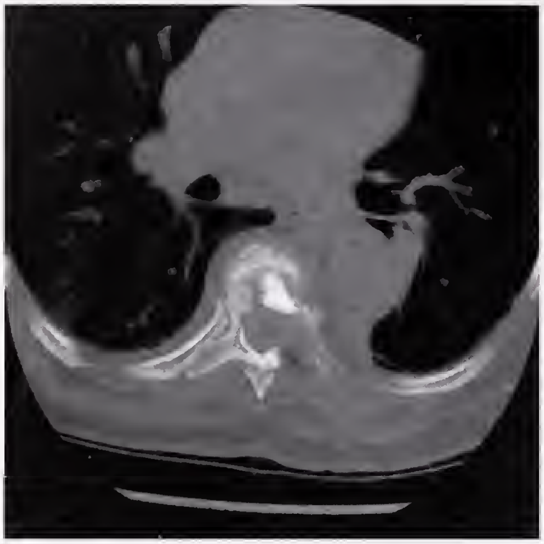


Fig 3. Postoperative computed tomography demonstrating minimally invasive transpedicular decompression with methyl methacrylate reconstruction of the vertebral body.

improvements in neurological deficits and ambulation status (Figure 3). In addition, a transpedicular injection of PMMA with or without the use of a balloon, kyphoplasty, and vertebroplasty can ameliorate pain in patients with spinal metastases.[21] Future directions may include the use of chemotherapy-impregnated cement for improved local control of disease.[22]

CONCLUSIONS

Spinal metastases are common in patients with various types of systemic cancers and can significantly impair the quality of life of patients by leading to pain and paralysis. Surgery should play an active role in a multidisciplinary approach to the management of many patients with metastatic cancer. Approaches and treatment options can be tailored to the individual's pathology, life expectancy, and symptoms to maximize palliation, tumor control, and quality of life.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Constans JP, de Divitiis E, Donzelli R, et al. Spinal metastases with neurological manifestations. Review of 600 cases. *J Neurosurg* 1983; 59: 111–118.
2. Sundaresan N, Boriani S, Rothman A, Holtzman R. Tumors of the osseous spine. *J Neurooncol* 2004; 69: 273–290.
3. Gasbarrini A, Cappuccio M, Mirabile L, et al. Spinal metastases: treatment evaluation algorithm. *Eur Rev Med Pharmacol Sci* 2004; 8: 265–274.
4. Wai EK, Finkelstein JA, Tangente RP, et al. Quality of life in surgical treatment of metastatic spine disease. *Spine* 2003; 28: 508–512.
5. Li KC, Poon PY. Sensitivity and specificity of MRI in detecting malignant spinal cord compression and in distinguishing malignant from benign compression fractures of vertebrae. *Magn Reson Imaging* 1988; 6: 547–556.
6. Gilbert RW, Kim JH, Posner JB. Epidural spinal cord compression from metastatic tumor: diagnosis and treatment. *Ann Neurol* 1978; 3: 40–51.
7. Greenberg HS, Kim JH, Posner JB. Epidural spinal cord compression from metastatic tumor: results with a new treatment protocol. *Ann Neurol* 1980; 8: 361–366.
8. Rodriguez M, Dinapoli RP. Spinal cord compression: with special reference to metastatic epidural tumors. *Mayo Clin Proc* 1980; 55: 442–448.
9. Sorensen S, Borgesen SE, Rohde K, et al. Metastatic epidural spinal cord compression. Results of treatment and survival. *Cancer* 1990; 65: 1502–1508.
10. Young RF, Post EM, King GA. Treatment of spinal epidural metastases. Randomized prospective comparison of laminectomy and radiotherapy. *J Neurosurg* 1980; 53: 741–748.
11. Algra PR, Heimans JJ, Valk J, et al. Do metastases in vertebrae begin in the body or the pedicles? Imaging study in 45 patients. *AJR Am J Roentgenol* 1992; 158: 1275–1279.
12. Sundaresan N, Rothman A, Manhart K, Kelliher K. Surgery for solitary metastases of the spine: rationale and results of treatment. *Spine* 2002; 27: 1802–1806.
13. Tomita K, Kawahara N, Murakami H, Demura S. Total en bloc spondylectomy for spinal tumors: improvement of the technique and its associated basic background. *J Orthop Sci* 2006; 11: 3–12.
14. Bach F, Larsen BH, Rohde K, et al. Metastatic spinal cord compression. Occurrence, symptoms, clinical presentations and prognosis in 398 patients with spinal cord compression. *Acta Neurochir (Wien)* 1990; 107: 37–43.
15. Cole JS, Patchell RA. Metastatic epidural spinal cord compression. *Lancet Neurol* 2008; 7: 459–466.
16. Schiff D, O'Neill BP, Suman VJ. Spinal epidural metastasis as the initial manifestation of malignancy: clinical features and diagnostic approach. *Neurology* 1997; 49: 452–456.
17. Helweg-Larsen S, Sorensen PS. Symptoms and signs in metastatic spinal cord compression: a study of progression from first symptom until diagnosis in 153 patients. *Eur J Cancer* 1994; 30A: 396–398.
18. Patchell RA, Tibbs PA, Regine WF, et al. Direct decompressive surgical resection in the treatment of spinal cord compression caused by metastatic cancer: a randomised trial. *Lancet* 2005; 366: 643–648.
19. Chi JH, Gokaslan Z, McCormick P, et al. Selecting treatment for patients with malignant epidural spinal

cord compression–does age matter?: results from a randomized clinical trial. *Spine* 2009; 34: 431–435.

20. Fourney DR, Gokaslan ZL. Spinal instability and deformity due to neoplastic conditions. *Neurosurg Focus* 2003; 14: e8.
21. Pflugmacher R, Taylor R, Agarwal A, et al. Balloon kyphoplasty in the treatment of metastatic disease of the spine: a 2-year prospective evaluation. *Eur Spine J* 2008; 17: 1042–1048.
22. Abe T, Sakane M, Ikoma T, et al. Intraosseous delivery of paclitaxel-loaded hydroxyapatitealginate composite beads delaying paralysis caused by metastatic spine cancer in rats. *J Neurosurg Spine* 2008; 9: 502–510.

NOTHING BEATS
BEING FIRST

SIGN UP FOR FREE
Journal of Hospital Medicine e-alerts
and be the first to know about new
research in your field!

It's fast, free, and flexible.
Customize your alerts to deliver
only the information you need.

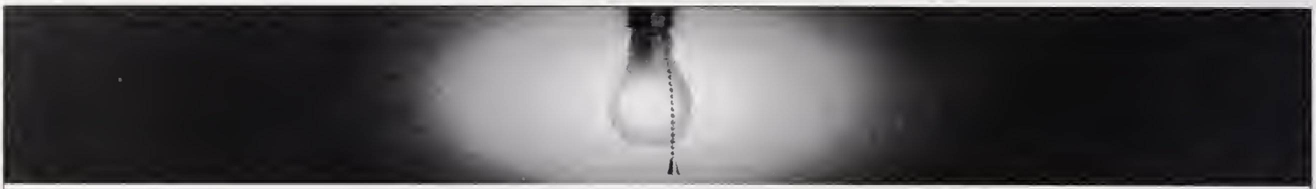
WILEY
InterScience®
DISCOVER SOMETHING GREAT

www.JournalofHospitalMedicine.com

SIGN UP TODAY!

WILEY-BLACKWELL

shm
Society of Hospital Medicine



don't publish online in the dark...

let our open linking network
shed some light.

This journal is CrossRef-ed! As a scientist or scholar navigating digital content, you've come to expect dynamic linking from citation to original source. As an author, you expect ready dissemination of your work to others. CrossRef is a collaborative, cross-publisher reference linking network that allows users to click on a citation and be taken directly to the full-text location. An independent, non-profit organization that employs open standards, CrossRef was established in cooperation with leading research publishers and the International DOI Foundation. Since the service went live in 2000, CrossRef has linked millions of articles from thousands of journals. Find out more at **www.crossref.org**, and make sure all your electronically published work is CrossRef-enabled.

crossref

CROSSREF.ORG | THE CITATION LINKING BACKBONE

200 WHEELER ROAD, 6TH FLOOR, BURLINGTON, MA 01803 • 781-359-2442

VOL. 77, NO. 2
MARCH/APRIL 2010

MOUNT SINAI JOURNAL OF MEDICINE

Journal of Translational and Personalized Medicine

Issue Theme
Genes, Race, Ethnicity and the Environment

Theme Editors
Gary Butts, MD
Kurt Hirschhorn, MD

RECEIVE
PR 1 4

ce, Ethnicity, Ancestry, and Pharmacogenetics
vironmental Justice and Children's Health
There a Genetic Basis for Health Disparities in HIV Disease? Yes and No
enetics and Susceptibility to Asthma in the U.S. Latino Population
enetic and Environmental Contributions to Racial Disparities in Preterm Birth
stice and Protection against Discrimination in the Use of Genetic Information
w Race and Ethnicity Impact Access to Health and Healthcare

pecial Features

lorectal Cancer Screening: Colonoscopy vs CT Colonography
igenetic Influences in Pregnancy and Early Development
pe 2 Diabetes and Cancer: What's the Connection?

WILEY-BLACKWELL

Recommend this Journal
to Your Library Today!

MOUNT SINAI JOURNAL OF MEDICINE

A Journal of Translational and Personalized Medicine

Editor-in-Chief: **Penny A. Asbell, MD, FACS, MBA**
Mount Sinai School of Medicine, New York, NY

Continuously published since 1934 and included in the *Index Medicus* from the very first volume, the **Mount Sinai Journal of Medicine (MSJM)** has a distinguished career as a general medical journal for a worldwide audience.

Dedicated to improving communication among all participants in medicine today, **MSJM** is an excellent source of up-to-date medical information for healthcare clinicians, educators, researchers and policy makers worldwide. **MSJM** will inspire the creative thinking needed to create new ways to understand, prevent, and treat human disease.

Each issue includes several quality, in-depth review articles related to a single theme and several others on subjects of diverse clinical interest. All theme articles are written by renowned authorities in their fields, from both Mount Sinai and elsewhere, under the supervision of Theme Editors.

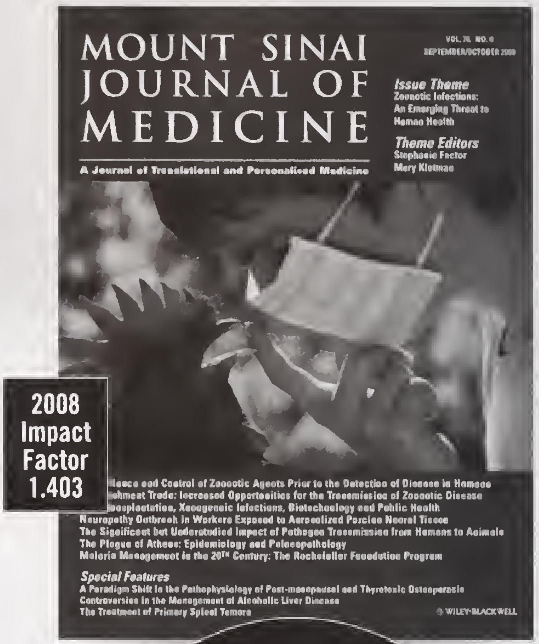
Upcoming Themes for Volume 77:

- Alzheimer's Disease: New Light on Etiology and Management
- Genes, Race, Ethnicity and the Environment
- Endovascular Surgery: Advances in Science and Technique
- Drug Discovery in Academic Settings
- The Obesity Pandemic
- Children's Health and the Environment

ACCESS THESE 2009 EDITORS CHOICE ARTICLES FREE ONLINE

- Advances and Controversies in Assisted Reproductive Technology (76:6)
- Xenotransplantation, Xenogenic Infection, Biotechnology and Public Health (76:5)
- The Student-run Health Clinic: A Novel Arena to Educate Medical Students on Systems-based Practice (76:4)
- Glycemic Targets for Patients with Type 2 Diabetes Mellitus (76:3)
- Neurobehavioral Sequelae of Traumatic Brain Injury (76:2)
- Spectrum of Fibrosing Diffuse Parenchymal Lung Disease (76:1)

Want more? Access a free sample issue online today!



MOUNT SINAI JOURNAL OF MEDICINE

A Journal of Translational and Personalized Medicine

VOL. 76, NO. 6
SEPTEMBER/OCTOBER 2009

Issue Theme
Zoonotic Infections:
An Emerging Threat to
Human Health

Theme Editors
Stephanie Factor
Mary Kleinman

llance and Control of Zoonotic Agents Prior to the Detection of Disease in Humans
shmeat Trade: Increased Opportunities for the Transmission of Zoonotic Disease
otransplantation, Xenogenic Infections, Biotechnology and Public Health
Neuropathy Outbreak in Workers Exposed to Aerosolized Porcine Neural Tissue
The Significant but Understudied Impact of Pathogen Transmission from Humans to Animals
The Plague of Athens: Epidemiology and Paleopathology
Malaria Management in the 20TH Century: The Rockefeller Foundation Program

Special Features
A Paradigm Shift in the Pathophysiology of Post-menopausal and Thyrotoxic Osteoporosis
Controversies in the Management of Alcoholic Liver Disease
The Treatment of Primary Splenic Tumors

WILEY-BLACKWELL

2008
Impact
Factor
1.403

Stay on Top of the Research!

Sign-up for Email Table of Contents Alerts delivered direct to your desktop – FREE!

Simply go to the homepage and click on "Set E-Mail Alert" for automatic table of contents alerts, or create a customized alert based on your own search criteria.

How to Receive Complimentary Online Access for Institutions?

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Library Recommendation Form

To________________________________

Date______________________________

Requested by:_______________________

Dept.______________________________

I have reviewed the **Mount Sinai Journal of Medicine** and urge you to add a complimentary 2010 online subscription to our subscription list. Please contact me should you have any questions.

WWW.MOUNTSINAIJOURNAL.ORG

WILEY-BLACKWELL

Enter Your 2010 Subscription Today!

MOUNT SINAI JOURNAL OF MEDICINE
Issue Theme
Type 2 Diabetes Mellitus: New Evidence, More Controversy
Theme Editor
Special Features

med at clinicians and educators, the ount Sinai Journal of Medicine provides jh-quality, in-depth review articles in ery issue, studded with clinical pearls.

ur 2010 will include the six 2010 print ues of the **Mount Sinai Journal of edicine** (Volume 77), plus concurrent f l-text online access to these six issues Wiley's Internet publishing platform.

andard Personal Subscriptions

$30.00 (U.S./Canada/Mexico)
$20.00 (Outside North America)

ount Sinai Alumni Subscriptions!

$.00 (Worldwide)

tended for Institutions in 2010

mplimentary line Access!

ur institution can opt to receive nplimentary online access to the **unt Sinai Journal of Medicine** during 2 0. When registering for complimentary o ine access, your institution can also o ain one complimentary print s scription.

PERSONAL SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 personal subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $105.00 (U.S./Canada/Mexico) ☐ $130.00 (Outside North America)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

Method of Payment (Canada residents please add GST. Personal rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **John Wiley & Sons, Inc.** Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA

Card #________________ Exp. Date________

Signature____________________
(credit card orders not valid without signature)

Ship To:

Name____________________

Address____________________

City____________ State______ Zip______

MSJM0910

MSJM ALUMNI SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 Alumna/Alumnus Subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $40.00 (Worldwide)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

METHOD OF PAYMENT (Alumna/Alumnus rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **The Mount Sinai School of Medicine**. Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA ☐ Other

Card #________________ Exp. Date________

Signature____________________
(credit card orders not valid without signature)

Ship To:

Name____________________

Address____________________

City____________ State______ Zip______

To subscribe online visit: **www.msjmonline.com**

MSJM10

COMPLIMENTARY ONLINE ACCESS FOR INSTITUTIONS

Mount Sinai Journal of Medicine

A Journal of Translational and Personalized Medicine

2010. Volume 77. 6 issues per year. Print ISSN: 0027-2507. Online ISSN: 1931-7581.

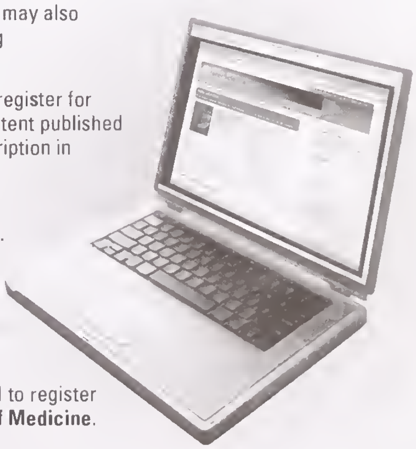
Your institution can opt to receive complimentary online access to the **Mount Sinai Journal of Medicine** during 2010. Your institution may also request one complimentary print subscription when registering for complimentary online access.

There is no obligation to subscribe in subsequent years if you register for complimentary access. Institutions retain online access to content published during the complimentary period, with or without a paid subscription in subsequent years.

Dear Librarian,
I have reviewed the scope of Mount Sinai Journal of Medicine.
Please register for complimentary access:

Name: ____________ Department: ____________

Signature: ____________ Date: ______



Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Don't delay, recommend to your library today!

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO. 2277 HOBOKEN, NJ

POSTAGE WILL BE PAID BY ADDRESSEE

JOHN WILEY & SONS INC
ATTN: BETH MOLD, MAIL STOP 8-01
111 RIVER STREET
HOBOKEN, NJ 07030-9441

PLACE
POSTAGE
STAMP
HERE

INES SANTOS
MOUNT SINAI JOURNAL OF MEDICINE
ONE GUSTAV L. LEVY PLACE BOX 1094
NEW YORK, NY 10029-6574

Library Routing Card

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

Editor-in-Chief

Penny A. Asbell

Deputy Editor

David H. Adams

Associate Editors

Stuart Aaronson
Erwin Bottinger
Dennis S. Charney
Valentin Fuster
Wayne Goodman
Jonathan Halperin
Ravi Iyengar
Paul Klotman
Ihor R. Lemischka
Michael L. Marin
Thomas G. McGinn
Ira Nash
Maria New
Rosamond Rhodes
Hugh Sampson
Mary Sano
Jeffrey H. Silverstein
Albert Siu
Anthony Squire
Fred Suchy
Savio Woo

Editorial Board

William J. Bremner	University of Washington
Thomas M. Coffman	Duke University
Dwight Evans	University of Pennsylvania
Gabriel Haddad	University of California, San Diego
Heidi E. Hamm	Vanderbilt University
Matti Häyry	University of Manchester
Eve J. Higginbotham	Howard University
Katherine A. High	University of Pennsylvania
Tony Hope	University of Oxford
Mary Klotman	Duke University
John H. Krystal	Yale University
Raju Kucherlapati	Harvard Medical School
Robert J. Lefkowitz	Duke University
Robert Lisak	Wayne State University
George Lister	University of Texas
Philip A. Marsden	University of Toronto
Charles Nemeroff	Emory University
Scott Rauch	Harvard Medical School
Bernard Roizman	University of Chicago
Alan F. Schatzberg	Stanford University
Inder Verma	The Salk Institute
Jeffrey A. Whitsett	University of Cincinnati
Allan Wolkoff	Albert Einstein College of Medicine
Noam Zohar	Bar-Ilan University

Theme Editors

Gary Butts, MD
Associate Dean for Diversity Programs
Policy and Community Affairs

Kurt Hirschhorn, MD
Professor Emeritus of Pediatrics, Genetics and
Medicine; Chairman Emeritus of Pediatrics

Mount Sinai School of Medicine, New York, NY

Editorial Director

Stephen Cohen
stephen.cohen@mssm.edu
212-241-6304

Publisher

John Wiley & Sons
Anna Wesley, Associate Editor
awesley@wiley.com

Genes, Race, Ethnicity and the Environment

Future *MSJM* Themes and Theme Editors

- Obesity (Robert Yanagiswa, Derek LeRoith)
- Drug discovery in academic settings (Ravi Iyengar)
- Biomarkers in clinical practice (penny Asbell, Mark Babyatsky)
- Antherosclerosis (David Adams, Valentin Fuster)
- Future of primary care (Thomas McGinn)
- Parkinson's disease (Michele Tagliati, Ron Alterman)
- Children's health and the environment (Philip Landrigan)
- Endovascular surgery (Michael Marin. Peter Faries)
- Translational approches to cancer treatment (Steven Burakoff, William Oh)

Error Correction

The January-February issue of the MSJM (Vol 77, No.1) incorrectly stated the academic affiliations of Allal Boutajangout, MD, co-author of an article entitled "Vaccination as a Therapeutic Approach to Alzheimer's Disease." Dr. Boutajangout's departmental affiliations at the New York University School of Medicine are the Department of Psychiatry and Department of Physiology and Neuroscience. The article also omitted to state that Dr. Boutajangout's contribution to the article was supported by grant MNIRGD-09-133917 from the Alzheimer's Association.

©**2010 Mount Sinai School of Medicine.** All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, Phone: (978) 750-8400; Fax: (978) 750-4770. Such requests and other permission inquiries should be addressed to the Permissions Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; Phone: (201) 748-6011; Fax: (201) 748-6008; E-mail: permreq@wiley.com. Printed in the United States of America by The Sheridan Press.

Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine (Print ISSN: 0027-2507; Online ISSN: 1931-7581 at Wiley InterScience, www.interscience.wiley.com) is published twelve times per year for the Mount Sinai School of Medicine by Wiley Subscription Services, a Wiley Company, 111 River Street, Hoboken, NJ 07030.

Send subscription inquiries to: Subscription Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-6645.

Offprint sales and inquiries should be addressed to the Reprint Billing Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-8776. All other inquiries should be addressed to the Customer Service Department at (201) 748-6645.

Individual subscription price: $95.00 in North America, $120.00 outside North America. Payment must be made in US dollars drawn on a US bank. Claims for undelivered copies will be accepted only after the following issue has been delivered. For more subscription prices and information, please visit www.mountsinaijournal.org.

Change of address: Please forward six weeks prior to move: present mailing label with change of address. Please enclose a copy of the mailing label or cite your subscriber reference number in order to expedite handling. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits. Send claims in care of: *Mount Sinai Journal of Medicine*, c/o John Wiley & Sons, Inc., Attn: Journals Admin Dept. UK, 111 River Street, Hoboken, NJ 07030.

Cancellations: Subscription cancellations will not be accepted after the first issue has been mailed. Periodicals postage paid at Hoboken, NJ and at additional mailing offices.

Postmaster: Send address changes to *Mount Sinai Journal of Medicine*, Subscription Distribution, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Printed in the United States of America. Single issues from current and recent volumes are available at the current single issue price from cs-journals@wiley.com. Earlier issues may be obtained from Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. Tel: +1 518 537 4700, Fax: +1 518 537 5899, Email: psc@periodicals.com

Advertising inquiries should be addressed to: Advertising Sales Manager, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Phone: (201) 748-6921; Fax: (201) 748-6207. Advertising in this journal is published with a caveat emptor understanding. The Publisher, Editors, and members of the Editorial Board do not imply warranty of products, nor quality, validity, or approval of education material offered by such advertising.

Indexed by: Biological Abstracts® (Thomson ISI), Chemical Abstracts Service/SciFinder (ACS), Current Contents®/Clinical Medicine (Thomson ISI), EMBASE/Excerpta Medica (Elsevier), Index Medicus/MEDLINE/PubMed (NLM), Science Citation Index Expanded™ (Thomson ISI), SCOPUS (Elsevier), Web of Science® (Thomson ISI).

This journal is printed on acid-free paper.

Disclaimer: All articles published, including but not limited to original research, reviews, clinical notes, clinical pearls, prismatic cases, clinician's corner, editorials, reports, and letters, represent the opinions and views of the authors and do not reflect any official policy or medical opinion of the Mount Sinai School of Medicine, the institutions with which the authors are affiliated, or the Publisher unless this is clearly specified. Articles published herein should not be relied upon as recommending or promoting a specific method, diagnosis, or treatment by physicians for any particular patient or disorder. Although the Editors and Publisher believe that drug selections and dosages and the specifications and usage of equipment and devices as set forth herein are in accord with current recommendations and practice at the time of publication, they accept no legal responsibility for any errors or omissions, and make no warranty, express or implied, with respect to material contained herein.

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

CONTENTS
VOL 77, NO 2

Theme Introduction

131 Genes, Race, Ethnicity and the Environment
Gary Butts, Kurt Hirschhorn

Theme Articles

133 Race, Ethnicity, Ancestry and Pharmacogenetics
Thomas J. Urban

140 Genetics and Asthma Disease Susceptibility in the U.S. Latino Population
Joan Reibman, Mengling Liu

149 Is There a Genetic Basis for Health Disparities in HIV Disease? Yes and No
Cheryl Winkler

160 Genetic and Environmental Contributions to Racial Disparities in Preterm Birth
Siobhan Dolan

166 Access to Health and Healthcare: How Race and Ethnicity Matter
Lynne D. Richardson, Marlaina Norris

178 Environmental Justice and the Health of Children
Philip J. Landrigan, Virginia A. Rauh, Maida P. Galvez

188 You're Born with Your Genes: Justice and Protection Against Discrimination in the Use of Genetic Information
Leslie P. Francis

Special Features

197 Type 2 Diabetes and Cancer: What's the Connection?
Dara Cannata, Yvonne Fierz, Archana Vijayakumar, Derek LeRoith

214 Colonoscopy vs Computerized Tomographic Colongraphy for Colorectal Cancer Screening
Reza Akhtar, Michelle Lee, Steven H. Itzkowitz

225 Epigenetic Influences in Pregnancy and Early Development
Yevgeniya Pozharny, Luca Lambertini, Garfield Clunie, Lauren Ferrara, Men-Jean Lee

SUBMIT YOUR WORK TO

Mount Sinai Journal of Medicine

The *Mount Sinai Journal of Medicine* state-of-the-art online submission system allows you to track your paper through the entire production process. You will never have to wait, or wonder, when your paper will publish.

Submission is fast, easy, and safe.
Follow these 5 simple steps:

LOG ON TO

http://mc.manuscriptcentral.com/msjm
Log on with your user ID and password. Click "Author Center" and then "Submit First Draft of a New Manuscript."

TITLE, ABSTRACT & KEYWORDS

Copy and paste these from your document. Use the "Character Palette" to add Greek and other symbols.

UPLOAD FILES USING THE FILE MANAGER

Use the "Browse" button to select a file. Choose a designation from the list and whether it will be seen by referees. Click "Upload" to add your file. Repeat until all files are uploaded.

VIEW 'PROOF'

You must check the 'proofs' before submission. This will show you what Editors and Reviewers will see. You can go back and replace incorrect files.

SUBMIT AND TRACK YOUR PAPER

Click "Submit" once you are satisfied. An automatic confirmation with your manuscript number will follow and the progress of your submission can be tracked.

SUBMISSION HELP

Online
Select the red "Get Help Now" link in the upper right-hand corner of the login screen.

Phone
434-964-4100 (local)
888-503-1050
(domestic toll-free)

REQUIREMENTS

File Types
All types can be uploaded but not all users will have access to all software. To minimize problems we recommend you upload text as Word or RTF, and figures as JPEG, GIF, TIFF, BMP, PNG, EPS, Powerpoint or Excel.

Where possible combine multiple files into a single document.

Submit at:

http://mc.manuscriptcentral.com/msjm

WILEY-BLACKWELL

THEME INTRODUCTION

Genes, Race, Ethnicity, and Environment

Gary C. Butts, MD and Kurt Hirschhorn, MD

Pediatrics, Medical Education, and Preventive Medicine, Mount Sinai School of Medicine,
New York, NY

The Center for Multicultural and Community Affairs at Mount Sinai School of Medicine aims (1) to support the elimination of health disparities and inequities through a nationally recognized center structure using innovative, integrative, and coordinated approaches to the areas of community affairs, patient care, education, and research; (2) to improve the health of all populations through diversification of the health care workforce; and (3) to influence health policy and research. As such, we convene meetings and conferences to present relevant scientific updates and facilitate discussion on issues and concerns particularly related to improving health outcomes for minority communities and vulnerable populations.

In November 2008, the Center for Multicultural and Community Affairs sponsored a symposium on genes, race, ethnicity, and the environment at Mount Sinai with the following aims: (1) to provide an update on scientific insights into gene expression and disease prevalence and variation; (2) to provide background information on how race, ethnicity, and the environment affect health; and (3) to develop a dialogue on health equity in light of the National Institutes of Health vision of transforming medicine and health through science. The symposium presented scientific updates and conversations among academics and community stakeholders as it has become increasingly clear that coupling advances in genetics and genomic science to the health disparities agenda is critical to understanding the full array of complex contributors that need to be further elucidated and subsequently addressed to adequately reduce disparities and inequities in health and health care.

The articles included in this theme issue of the *Mount Sinai Journal of Medicine* have been provided by many of the presenters from the symposium and, in large part, reflect the scientific content presented. The authors highlight recent advances in genetics and genomics with respect to the risks and management of specific diseases and conditions such as human immunodeficiency virus, asthma, and preterm birth, and they note how persistent concerns about access to quality health care and a disproportionate burden of environmental hazards on the poor and racial and ethnic minorities present a social justice challenge for addressing health care disparities. In addition, the authors suggest unique ethical challenges that we must address in balancing the value of access to genetic information for personalized management with the potentially added risk of sharing information for individuals, particularly from vulnerable populations. Arguably, there are few areas in health care and research challenged with controversy like that posed by the subject of genes and race.

In *Race, Ethnicity, Ancestry, and Pharmacogenetics*, Thomas J. Urban comments on the benefits of pharmacogenetics in supporting efforts to target groups on the basis of differences in drug response, but he also suggests that these advances provide hope of moving to individualized/personalized drug prescribing, which will minimize the need for race, ethnicity, and ancestry as proxies for differentiating management.

Cheryl Winkler, in *Is There a Genetic Basis for Health Disparities in Human Immunodeficiency Virus Disease? Yes and No*, notes that although research does not consistently support a genetic predisposition of susceptibility to human immunodeficiency virus infection for individuals of African descent, the data do support the idea that increased risk to human immunodeficiency virus–associated nephropathy is due in part to susceptibility genes

Address Correspondence to:

Gary C. Butts
Pediatrics, Medical Education,
and Preventive Medicine
Mount Sinai School of Medicine
New York, NY
Email: gary.butts@mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20177

© 2010 Mount Sinai School of Medicine

that are very frequent throughout sub-Saharan Africa but infrequent or absent in non-Africans. However, she also notes that it is essential that well-designed studies be carried out in diverse populations.

In *Genetics and Asthma Disease Susceptibility in the US Latino Population*, Joan Reibman and Mengling Liu posit that the high prevalence of asthma in specific Latino subpopulations makes it imperative to study asthma in Hispanics; however, the study of asthma and asthma genetics is made complex in these populations by the diversity and heterogeneity of the Hispanic population in the United States. These studies have the potential to enhance risk identification, target pharmacological therapy, and guide environmental regulations and thus to reduce health disparities.

Siobhan M. Dolan, in *Genetic and Environmental Contributions to Racial Disparities in Preterm Birth*, stresses that efforts to identify relevant gene-environment interactions affecting contributors to preterm birth may support research and targeted strategies to reduce persistent racial and ethnic disparities in preterm birth and infant mortality in the United States.

In *Environmental Justice and the Health of Children*, Philip J. Landrigan, Virginia A. Rauh, and Maida P. Galvez point out that the inequitable and disproportionately heavy exposure of poor, minority, and disenfranchised populations to toxic chemicals and other environmental hazards contributes to disparities in health status across populations of differing ethnicity, race, and socioeconomic status and that documentation of linkages between health disparities and environmental injustice is an important step toward achieving environmental justice.

Lynne Richardson and Marlaina Norris, in *Access to Health and Health Care: How Race and Ethnicity Matter*, remind us that racial and ethnic disparities in health are multifactorial and reflect differences in biological vulnerability to disease as well as differences in social resources, environmental factors, and health care interventions. The authors note that as genomics helps us to elucidate the mechanisms by which genetic factors influence health and by which genetic susceptibility interacts with behavioral, nutritional, pharmacological, and environmental variables, it is also essential to understand how differential exposures to various behavioral, nutritional, pharmacological, and environmental variables contribute to health disparities.

Finally, in *You Are Born with Your Genes: Justice and Protection Against Discrimination in the Use of Genetic Information*, Leslie P. Francis describes what she calls the paradox of genetic information for justice in health care. On the one hand, it seems unjust to take into account genetic information in the distribution of important goods, but on the other hand, genetic information is useful for increasing personalized medicine. She notes that historical beliefs based on race and racial discrimination pose significant concerns, and it is important to protect the use of genetic information, particularly for vulnerable populations.

It appears that the pace of new knowledge available from genetic and genomic research, which will inform the health disparities discussion and therapeutic decisions, will continue to increase exponentially. Historically, new discoveries have not always directly benefited all communities equally, so it is important not to miss the opportunity to translate new advances to targeted groups and individuals who might benefit the most on the basis of risk and susceptibility. Additionally, to maximize the validity of this research and the potential to generalize its results, it is important to include participants from broadly diverse racial and ethnic groups and subpopulations. Providing a rationale that demonstrates concrete benefits that overcome the persistent historical mistrust by minority and vulnerable populations is critical.

We hope that the presented articles will prove useful both for supporting additional research to help further elucidate the relationships between genes and the environment and the complex interactions that contribute to disparities in health outcomes and for stimulating a broader conversation that includes social justice, ethics, and health care access.

ACKNOWLEDGMENTS

We offer special thanks to the editors of the *Mount Sinai Journal of Medicine* for supporting the publication and timely presentation of this theme issue. Furthermore, we must thank and acknowledge the expert advice, support, and input of the symposium planning committee, which was chaired by Kurt Hirschhorn and included Robert Desnick, Rosamond Rhodes, and Kimberly Morland, all of whom also supported the preparation of this theme issue.

DISCLOSURES

Potential conflict of interest: Nothing to report.

Race, Ethnicity, Ancestry, and Pharmacogenetics

Thomas J. Urban, PharmD, PhD

Duke University, Durham, NC

OUTLINE

Race and Pharmacogenetics
Race, Ethnicity, and Ancestry
Genome-Wide Association Studies and Health Disparities
Pharmacogenetics in the Future: Pharmacogenomics and Rare Variation

ABSTRACT

Pharmacogenetics is the study of how genetic variation influences the response to drugs. The concepts of race, ethnicity, and ancestry have long had a strong influence on pharmacogenetic discovery and on our understanding of population-level differences in drug response. The primary goal of pharmacogenetics, however, is to identify the individual genetic determinants of drug activity so that therapy can be tailored to the individual patient. This article describes the relationship between the concepts of race, ethnicity, and ancestry and how these concepts have been applied to pharmacogenetics, and it provides examples of the benefits and pitfalls associated with the use of racial or ethnic labels in genetic studies. The future of pharmacogenetics, including the study of rare genetic variation and what this means for racial or ethnic disparities in pharmacogenetic discovery, is also discussed. ***Mt Sinai J Med 77:133–139, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: ethnicity, genome-wide association studies, pharmacogenetics, race.

Address Correspondence to:

Thomas J. Urban
Duke University
Durham, NC
Email: thomas.urban@duke.edu

Pharmacogenetics, the field of study devoted to identifying genetic variants that predict the response to drugs, began in earnest over 50 years ago with the discovery that the metabolism of certain drugs, such as the muscle relaxant succinylcholine and the antituberculosis drug isoniazid, has a genetic component.[1] Since that time, a tremendous amount of effort has been expended, and much knowledge has been gained, with respect to the individual genes and particular variants of those genes that lead to higher or lower rates of therapeutic response to drugs or to drug toxicity.[2] For the most part, the field has focused on genetic determinants of pharmacokinetics, that is, the genes involved in drug absorption, metabolism, and transport across epithelial barriers such as the bile canalicular membrane (for biliary excretion) and the renal tubule epithelium (for urinary excretion). The major focus on pharmacokinetics is understandable: as most drugs have a predictable relationship between the concentration in blood or plasma and their effects, whether therapeutic or toxic, an individual who is likely to clear a particular drug at a lower-than-average rate is expected to experience correspondingly higher blood levels, and thus lower-than-standard doses of the drug may be required for that individual (and vice versa). Knowledge of the genetic basis of drug metabolism/elimination can thus, in theory, translate readily to clinical decision making.

However, there are also a growing number of examples, some of which are described later, of genetic determinants of pharmacodynamics, that is, how some drugs bring about their therapeutic effects (eg, through receptor signaling) or adverse effects (eg, through the elicitation of a maladaptive immune response). As we continue to march into the genomic era, it is anticipated that many additional genetic predictors of drug response involving complex and unknown interactions with many and varied biochemical targets will be identified.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20168

© 2010 Mount Sinai School of Medicine

As we continue to march into the genomic era, it is anticipated that many additional genetic predictors of drug response involving complex and unknown interactions with many and varied biochemical targets will be identified.

Thus, the study of pharmacogenetics can be described as having 2 main goals. The primary and more ambitious goal is to use individual genetic information to inform drug prescribing (eg, the choice of medication, dose, and treatment duration). The second and more easily achievable goal is to identify genes and gene pathways that are involved in drug response in order to gain biological knowledge about the activity of a drug and potentially about the pathophysiology underlying the disease that the drug is used to treat. Thus, even in cases in which pharmacogenetics does not tell us exactly what dose of what drug to prescribe to what person, it may inform our overall understanding of a disease and its treatment and lead potentially to more rational avenues for future drug development.

Even in cases in which pharmacogenetics does not tell us exactly what dose of what drug to prescribe to what person, it may inform our overall understanding of a disease and its treatment and lead potentially to more rational avenues for future drug development.

RACE AND PHARMACOGENETICS

There has for some time been a clinical suspicion that the concepts of race and ethnicity matter somehow in terms of outcomes after drug treatment, although these ethnic disparities and their underlying causes with respect to therapeutic outcome have often not been well studied empirically. Perhaps the most well-studied examples come from the field of cardiovascular disease,[3] in which as early as 1982 empirical evidence suggested that the response to antihypertensive drugs differed by race, with individuals who identified themselves as white having a higher response to beta-adrenergic receptor antagonists (ie, beta-blockers) than individuals who identified themselves as black.[4] Such a difference could be explained by any number of environmental or genetic factors as well as interactions between them, but one part of the explanation may be that there are common genetic variants that differ in frequency between black and white populations and that these variants control the response to beta-blockers. In the years since, we have witnessed an accumulation of evidence for this hypothesis: multiple single nucleotide polymorphisms have been discovered in the genes encoding the beta-adrenergic receptors, which are the targets of this class of drugs, and these single nucleotide polymorphisms have been shown or predicted to affect either the function of the receptors or their interactions with beta-blockers.[5] The same variants often show marked differences in frequency among world populations, such that individuals of primarily European ancestry are more likely to carry the high-response alleles at these sites in comparison with individuals of primarily African ancestry. Although these findings may be entirely coincidental, this initial example at least provides support for the idea that common genetic variation can contribute to differences in drug response by self-identified ethnicity or race.

Another example is provided by the anticoagulant drug warfarin, for which it has been shown that the average dose required to maintain therapeutic effect differs substantially among ethnicities, with Asians requiring much lower doses than African Americans and Caucasians and Hispanics having intermediate dose requirements.[3] In this example, it is much more certain that population differences in the frequency of genetic variants in genes controlling warfarin metabolism [cytochrome P450 2C9 (*CYP2C9*)/cytochrome P450 2C19] and activity [vitamin K epoxide reductase complex subunit 1 (*VKORC1*)] are responsible for the ethnic differences in dose requirements.[3,6] Figure 1 shows the geographic distribution of one variant controlling the warfarin maintenance dose, a variant in the enzyme CYP2C9 that reduces enzymatic activity, with the low-dose-requiring allele highlighted in red. Currently, the Food and Drug Administration recommends that clinicians consider testing for the *CYP2C9* and *VKORC1* genotype before prescribing warfarin; however, the clinical importance of genetic testing for warfarin dosing remains controversial because of questions of cost versus benefit.[7,8]

Occasionally, the correlation between ethnicity and drug response, even when the mechanism is

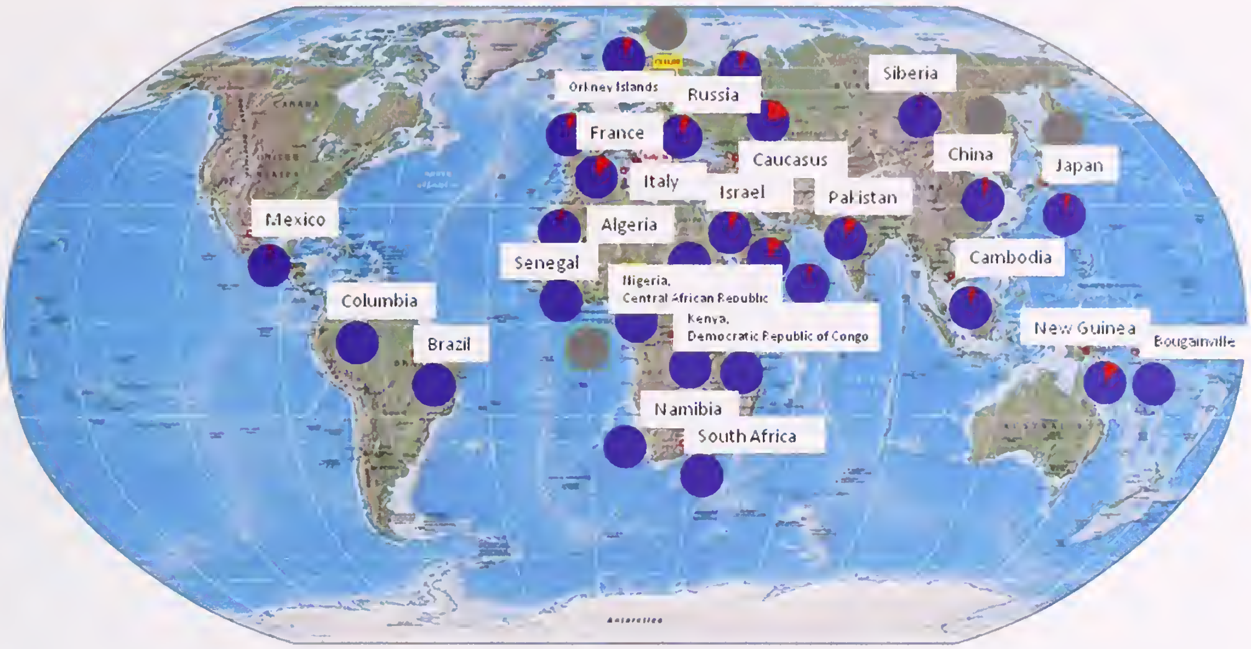


Fig 1. Distribution of the *CYP2C9*3* allele (shown in red in the pie charts) in various world populations. This variant results in a reduced capacity for warfarin metabolism and, therefore, lower dose requirements in *CYP2C9*3* carriers. The geographic distribution of *CYP2C9*3* along with other warfarin-sensitive alleles correlates with known ethnic differences in warfarin dose requirements.[3,6] Genotype data were generated by the Human Genome Diversity Project/Center for the Study of Human Polymorphism panel.[28] The illustration was generated with Diversity Viewer software, which was developed by Professor Dongliang Ge (Center for Human Genome Variation, Duke Institute for Genome Sciences and Policy). **Abbreviation:** CYP2C9, cytochrome P450 2C9.

unknown, is enough to encourage regulators to recommend that clinicians consider ethnicity per se in their prescribing practices. The emblematic example is the drug BiDil (hydralazine plus isosorbide dinitrate), which originally failed to show benefit as an antihypertensive drug in an ethnically mixed sample[9,10] but was later shown to be efficacious in African Americans in particular.[11,12] BiDil thus became the first drug to be approved specifically for treatment of a common disease in a particular ethnic group. Similarly, rosuvastatin has been shown to have substantially lower clearance and thus lower dose requirements in Asian populations versus European populations.[13] For both of these examples, there is the possibility that a common genetic variant with strong frequency differences between populations can explain the average population differences in the response observed in clinical studies. Such variants have not been identified for these examples but may be in the future as contemporary genetic approaches are applied to these and other problems in the prediction of drug response.

Importantly, the treatment of ethnicity as a proxy for ancestry, which is itself being used implicitly as a proxy for genotype, is a poor method for predicting treatment outcome. Again, the primary goal of pharmacogenetics is to use an individual's genetic information to improve drug outcomes. Although average ethnic differences in drug response may in some cases help us to find the individual genetic determinants of the variation in drug response, ultimately individual-level information is the key to determining with highest certainty how an individual

Importantly, the treatment of ethnicity as a proxy for ancestry, which is itself being used implicitly as a proxy for genotype, is a poor method for predicting treatment outcome. Although average ethnic differences in drug response may in some cases help us to find the individual genetic determinants of the variation in drug response, ultimately individual-level information is the key to determining with highest certainty how an individual patient will respond to treatment.

patient will respond to treatment. For example, there exist African Americans who receive no benefit from BiDil, and there are likely many individuals from non-African populations that might receive particular benefit from the same drug. If there is a genetic basis for the population-level differences in treatment response with BiDil, knowledge of a patient's genotype at the particular genomic site(s) influencing response will be much more useful than race or ethnicity as an indicator for prescribing this drug.

RACE, ETHNICITY, AND ANCESTRY

A point of clarification may be necessary in order to avoid conflation of several different concepts: race, ethnicity, and ancestry. From the perspective of population genetics, consideration of ancestry (ie, the structure of the relationship, however distant, among nominally unrelated individuals) in a population-based study is essential to avoid false-positive genetic findings due to confounding. Many gene variants differ in frequency between populations or subpopulations, but this is often merely due to random fluctuations (called genetic drift) and has no true biological meaning. If there are differences in risk for disease (or ascertainment) between populations in samples collected for a genetic association study, any and all of these neutral variants will thus give a signal of association that is biologically meaningless. Thus, failing to take account of ancestral structure can lead to spurious associations between diseases (or drug responses) and gene variants. For example, a variation in the lactase-encoding gene *LCT* has been associated with adult height in unmatched samples from a variety of European ancestral sources. However, this association has been shown to be spurious as both height and lactose persistence are correlated among European individuals at a population-wide level: matching subjects on the basis of the particular region of European origin greatly reduced the apparent association between *LCT* variants and adult height.[14] Contemporary genetics is now, however, equipped with the proper tools for managing this problem of confounding due to population structure.[15]

The terms *race* and *ethnicity*, on the other hand, are not biological terms but rather are sociological terms that describe group identification in discrete categories (with *ethnicity* having more currency, given its incorporation of cultural aspects of human groups, in contrast to *race*, which implies that humans can or should be categorized into discrete terms merely by genetics or biology).[16–18] Although the human phylogenetic tree has many branches and genetic ancestry often correlates (however weakly) with these sociological labels, a graded pattern of relatedness can be observed among individuals across the human diaspora. For several reasons, the categories of self-identified ethnicity are often used as a shorthand in genetic studies as a way of cutting the branches of the human species into reasonable subgroups for analysis. The primary reason is to maintain adequate statistical power for discovery: although the tools exist to correct for population structure even in combined ethnically diverse samples, doing so can potentially lead investigators to miss population-specific genetic effects or variants with consistent effects across populations when the allele frequency differs significantly among populations. An additional concern is that social and/or environmental correlates of ethnicity may modify true genetic effects. However, this shorthand should not be interpreted to mean that there is a biological basis for discrete categorization of humans into individual groups. Fortunately, humans as a species are far more complex and multifaceted than that, both genetically and sociologically.

GENOME-WIDE ASSOCIATION STUDIES AND HEALTH DISPARITIES

The current era of genome-wide association studies (GWASs) has been a revolution in human genetics and has brought with it a sobering sense of statistical rigor and an appreciation for the pitfalls associated with the previous era of human genetics (such as the previously described problem of confounding due to population structure). Although GWASs have now produced thousands of secure genetic associations

> *The current era of genome-wide association studies (GWASs) has been a revolution in human genetics and has brought with it a sobering sense of statistical rigor and an appreciation for the pitfalls associated with the previous era of human genetics.*

for many common human diseases, these studies have almost uniformly been disappointments in the sense that common genetic variation appears to contribute very little to disease risk.[19] This fact has led some to argue that rare genetic variants are likely to contribute much more substantially to disease risk,[19] whereas others remain hopeful that adequately powered GWASs may turn up additional common genetic risk factors and shed more light on the biology underlying common human diseases and other traits.[20]

For the most part, the GWASs that have been performed have been restricted to populations of primarily European descent, mostly because biomedical research funding is highest in the United States and Western Europe, where European ancestral populations make up the majority. Thus, a possible result of these studies might have been an increase in health disparities as the results of such studies might have revealed population-specific risk factors for diseases or drug responses, with the benefits of such research accruing only to the majority population. However, the lack of strong effects of common genetic variation in the GWASs that have been conducted in European populations to date suggests that, at least as some consolation, this preponderance of European-centric GWASs has not contributed to health disparities in terms of human disease genetics.[21]

> *The lack of strong effects of common genetic variation in the GWASs that have been conducted in European populations to date suggests that this preponderance of European-centric GWASs has not contributed to health disparities in terms of human disease genetics.*[21]

For other traits, however, there is more reason to believe that common genetic variation may play a significant role in outcomes, and this is most likely to be true for 2 arenas: infectious disease and drug response.[19] For example, very recently we identified via a GWAS a common genetic variant in the interleukin 28B (*IL28B*) gene, which encodes interferon lambda-3, and this is now the strongest known predictor of response to the standard treatment of chronic hepatitis C virus (HCV) infection (ie, pegylated interferon alfa plus ribavirin).[22] The same variant has also been shown to strongly influence the natural clearance of HCV.[23] The worldwide geographic distribution of the protective allele in this case is quite striking: it is almost fixed in East Asia (ie, almost all individuals are homozygous for the protective allele), whereas the frequency is lower in Europe and still lower in Africa. The reason for this geographic skew in allele frequencies may be due to differences in the survival advantage conferred by the protective allele in various parts of the world: some virus, perhaps now extinct, created a significant selection pressure especially in East Asia that drove the protective allele to fixation in that region of the world. Although the explanation for this geographic distribution of allele frequencies cannot be definitively ascribed to selection on the basis of currently available data, there exist examples of infectious diseases leading to observable evidence of selection in some populations, such as variants in the glycosyltransferase-like and dystrophin genes under selection pressure due to the Lassa virus.[24] This example serves as a caveat to the aforementioned general claim that common variants which differ in frequency among world populations are biologically neutral: within that set will be some variants that differ in frequency specifically because of differences in environmental pressures and population-specific natural selection. These may be infectious agents or small-molecule toxins, which might make infectious disease genetics and pharmacogenetics among the most fruitful areas in which to apply the GWAS strategy.

The *IL28B*/HCV story has particular importance for the idea of personalized medicine. Although it had previously been recognized that African Americans fare worse, on average, than European Americans in response to standard HCV treatment, the basis for this was previously unknown.[25] We now know that the *IL28B* genetic variant predisposing to poor treatment outcome is much more common in African American populations than European populations; in fact, this single variant alone explains roughly 50% of the average population differences in treatment response.[22] However, most importantly, an individual's genotype information trumps racial or ethnic labels: African Americans with the good-response genotype fared much better on average than individuals of European descent with the poor-responder genotype (Figure 2). This is the promise of pharmacogenetics: individual-level genetic information will inform treatment decisions in a manner on par with classical diagnostic tests and without the need for crude proxies such as race or ethnicity.

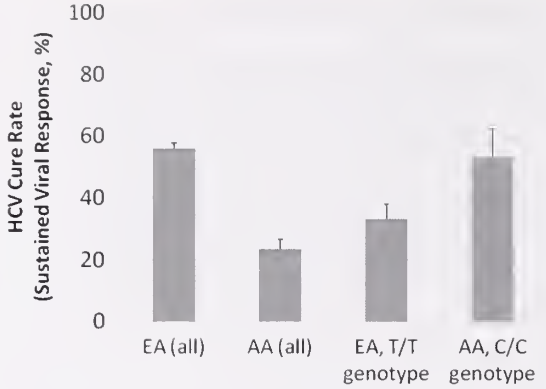


Fig 2. Sustained viral response to the standard treatment of HCV (pegylated interferon alfa plus ribavirin) as a function of global ancestry and the *IL28B* genotype. Although patients of primarily European descent tend to respond better on average than AA patients, the genotype at the treatment response–associated *IL28B* gene is a more informative predictor of response at the individual level. T/T corresponds to the low-response genotype at the associated marker (Single Nucleotide Polymorphism Database number rs12979860), whereas C/C corresponds to the high-response genotype. The data are reproduced from Ge *et al.*[22] **Abbreviations:** AA, African American; EA, European American; HCV, hepatitis C virus; IL28B, interleukin 28B.

> *This is the promise of pharmacogenetics: individual-level genetic information will inform treatment decisions in a manner on par with classical diagnostic tests and without the need for crude proxies such as race or ethnicity.*

PHARMACOGENETICS IN THE FUTURE: PHARMACOGENOMICS AND RARE VARIATION

Although the GWAS era is still current and will likely continue to add to our understanding of human disease and drug response, advances in human genetics will soon allow the comprehensive study of both common and rare genetic variation.[19] It has already been shown for several common diseases that rare, large-effect genetic variants explain much more of the risk for disease than common variants,[26] and this is likely to be true for at least some pharmacogenetic traits as well. For example, it has recently been shown that flucloxacillin-induced liver injury is associated with human leukocyte antigen B*5701, a low-frequency allele in a gene associated with antigen presentation and adaptive immunity.[27] Although this discovery was made with a GWAS approach, importantly, this risk variant lies just at the threshold allele frequency to have been discovered with this method. It is likely that many pharmacogenetic traits are also strongly influenced by rare variation, but their identification will require a comprehensive approach to rare variant identification, that is, whole genome or whole exome sequencing.

The approach to studying rare variation is likely to be quite different from that used for common variation.

> *Although the GWAS era is still current and will likely continue to add to our understanding of human disease and drug response, advances in human genetics will soon allow the comprehensive study of both common and rare genetic variation.*

Specifically, rather than using very large cohorts for identification (eg, n > 1000 cases in a typical GWAS), researchers will use a small number of carefully selected individuals exhibiting extreme values of the phenotype distribution for discovering rare genetic variants that explain the trait. The candidate variants arising from these discovery efforts may then be assessed for a causal role in the disease or drug response trait either by replication in larger cohorts of affected individuals or perhaps simply by comparison with a large reference panel of unaffected individuals. What this mean in terms of ethnic disparities in genetic studies is that self-identified ethnicity or genetic ancestry will begin to contribute much less to the decision to include or exclude research subjects from analysis: the availability of reference control genomes from diverse populations through projects such as the 1000 Genomes Project (http://www.1000genomes.org) promises to make the comparison of newly identified variants to the appropriate reference panel a simple matter in the next few years. Thus, it is hoped that the reliance on Western ethnic majority populations (ie, European ancestral populations) for genetic discovery will become a relic of the past and that the discovery of genetic variants associated with disease or drug response (and the biological and clinical relevance of these genetic effects) will be shared across ethnic boundaries.

> *It is hoped that the reliance on Western ethnic majority populations (ie, European ancestral populations) for genetic discovery will become a relic of the past and that the discovery of genetic variants associated with disease or drug response (and the biological and clinical relevance of these genetic effects) will be shared across ethnic boundaries.*

However, the social and environmental factors contributing to health disparities are an issue independent of any particular genetic approaches to studying human disease, and these will require continued attention.

DISCLOSURES

Potential conflict of interest: The author, Thomas J. Urban, is an inventor on a patent application related to IL28B variation in predicting HCV treatment response.

REFERENCES

1. Weinshilboum R. Inheritance and drug response. *N Engl J Med* 2003; 348: 529–537.
2. Evans WE, McLeod HL. Pharmacogenomics–drug disposition, drug targets, and side effects. *N Engl J Med* 2003; 348: 538–549.
3. Johnson JA. Ethnic differences in cardiovascular drug response: potential contribution of pharmacogenetics. *Circulation* 2008; 118: 1383–1393.
4. Veterans Administration Cooperative Study Group on Antihypertensive Agents. Comparison of propranolol and hydrochlorothiazide for the initial treatment of hypertension. II. Results of long-term therapy. *JAMA* 1982; 248: 2004–2011.
5. Muszkat M. Interethnic differences in drug response: the contribution of genetic variability in beta adrenergic receptor and cytochrome P4502C9. *Clin Pharmacol Ther* 2007; 82: 215–218.
6. Engen RM, Marsh S, Van Booven DJ, McLeod HL. Ethnic differences in pharmacogenetically relevant genes. *Curr Drug Targets* 2006; 7: 1641–1648.
7. Flockhart DA, O'Kane D, Williams MS, et al. Pharmacogenetic testing of CYP2C9 and VKORC1 alleles for warfarin. *Genet Med* 2008; 10: 139–150.
8. Kim MJ, Huang SM, Meyer UA, et al. A regulatory science perspective on warfarin therapy: a pharmacogenetic opportunity. *J Clin Pharmacol* 2009; 49: 138–146.
9. Cohn JN, Johnson G, Ziesche S, et al. A comparison of enalapril with hydralazine-isosorbide dinitrate in the treatment of chronic congestive heart failure. *N Engl J Med* 1991; 325: 303–310.
10. Cohn JN, Archibald DG, Ziesche S, et al. Effect of vasodilator therapy on mortality in chronic congestive heart failure. Results of a Veterans Administration cooperative study. *N Engl J Med* 1986; 314: 1547–1552.
11. Taylor AL, Ziesche S, Yancy C, et al. Combination of isosorbide dinitrate and hydralazine in blacks with heart failure. *N Engl J Med* 2004; 351: 2049–2057.
12. Carson P, Ziesche S, Johnson G, Cohn JN. Racial differences in response to therapy for heart failure: analysis of the vasodilator-heart failure trials. Vasodilator-Heart Failure Trial Study Group. *J Card Fail* 1999; 5: 178–187.
13. Lee E, Ryan S, Birmingham B, et al. Rosuvastatin pharmacokinetics and pharmacogenetics in white and Asian subjects residing in the same environment. *Clin Pharmacol Ther* 2005; 78: 330–341.
14. Campbell CD, Ogburn EL, Lunetta KL, et al. Demonstrating stratification in a European American population. *Nat Genet* 2005; 37: 868–872.
15. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; 38: 904–909.
16. Ledford H. Language: disputed definitions. *Nature* 2008; 455: 1023–1028.
17. Foster MW. Looking for race in all the wrong places: analyzing the lack of productivity in the ongoing debate about race and genetics. *Hum Genet* 2009; 126: 355–362.
18. Patten MA. 'Subspecies' and 'race' should not be used as synonyms. *Nature* 2009; 457: 147.
19. Goldstein DB. Common genetic variation and human traits. *N Engl J Med* 2009; 360: 1696–1698.
20. Hirschhorn JN. Genomewide association studies–illuminating biologic pathways. *N Engl J Med* 2009; 360: 1699–1701.
21. Need AC, Goldstein DB. Next generation disparities in human genomics: concerns and remedies. *Trends Genet* 2009; 25: 489–494.
22. Ge D, Fellay J, Thompson AJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 2009; 461: 399–401.
23. Thomas DL, Thio CL, Martin MP, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* 2009; 461: 798–801.
24. Sabeti PC, Varilly P, Fry B, et al. Genome-wide detection and characterization of positive selection in human populations. *Nature* 2007; 449: 913–918.
25. Muir AJ, Bornstein JD, Killenberg PG. Peginterferon alfa-2b and ribavirin for the treatment of chronic hepatitis C in blacks and non-Hispanic whites. *N Engl J Med* 2004; 350: 2265–2271.
26. Itsara A, Cooper GM, Baker C, et al. Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet* 2009; 84: 148–161.
27. Daly AK, Donaldson PT, Bhatnagar P, et al. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. *Nat Genet* 2009; 41: 816–819.
28. Cann HM, de Toma C, Cazes L, et al. A human genome diversity cell line panel. *Science* 2002; 296: 261–262.

Mount Sinai Journal of Medicine 77:140–148, 2010

Genetics and Asthma Disease Susceptibility in the US Latino Population

Joan Reibman, MD, and Mengling Liu, PhD

New York University School of Medicine, New York, NY

OUTLINE

Genetic Studies and Asthma Disparities
Complexity of Studying Genetics, Race and Ethnicity, and Health Disparities in the Latino Population
Genetics of Asthma
Selected Asthma Genes in the Latino Population
Conclusion

ABSTRACT

The US Latino population is heterogeneous with diversity in environmental exposures and socioeconomic status. Moreover, the US Hispanic population derives from numerous countries previously under Spanish rule, and many Hispanics have complex proportions of European, Native American, and African ancestry. Disparities in asthma severity and control are due to complex interactions between environmental exposures, socioeconomic factors, and genetic variations. In addition, diseases within the Latino community may also differ by country of origin. Although US Census data show low asthma rates in the Hispanic population as a whole, there is a lot of variability in the prevalence and morbidity of asthma, with a prevalence of 5.0% in Mexican Americans versus 17.0% in Puerto Ricans. The diversity and population admixture make the study of the genetics of asthma complex in Latino populations. However, an understanding of the genetics of asthma in all populations, including the Latino population, can enhance risk identification, help us to target pharmacological therapy, and guide environmental regulations, all of which can promote a reduction in health disparities. The inclusion of markers of ancestral diversity and the incorporation of techniques to adjust for stratification now make these studies feasible in complex populations, including the Latino population. To date, studies using linkage analyses, genome-wide associations, or candidate gene analyses have identified an association of asthma or asthma-related phenotypes with candidate genes, including interleukin 13, β-2 adrenergic receptor, a disintegrin and metalloproteinase 33, orosomucoid 1-like 3, and thymic stromal lymphopoietin. As reviewed here, although these genes have been identified in diverse populations, limited studies have been performed in Latino populations, and they have had variable replication. There is a need for the development of registries with well-phenotyped pediatric and adult Latino populations and subgroups for inclusion in the rapidly expanding field of genetic studies, and these studies need to be used to reduce health disparities. ***Mt Sinai J Med 77:140–148, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: asthma, genetics, Hispanic, Latino.

Address Correspondence to:

Joan Reibman
New York University School of Medicine
New York, NY
Email: joan.reibman@nyumc.org

Large-scale surveys conducted by the Centers for Disease Control and Prevention and the National Center for Health Statistics have estimated that 16.4 million adults (7.3%) and 7.0 million children (9.4%) in the United States currently have asthma, with a total lifetime asthma diagnosis prevalence of 7.7%.[1] These statistics vary geographically within the United States, with urban areas reporting higher rates of asthma, including a 12% self-reported lifetime prevalence of asthma among adults in New York City.[2] Although asthma is common in urban environments and is associated with socioeconomic status, the prevalence and severity of asthma also vary by race/ethnicity. In 2005–2007, the time period reported in the National Center for Health Statistics study, the rates of lifetime asthma diagnosis, current asthma, and asthma attacks were highest among American Indians/Alaska natives and blacks or African Americans in both

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20171

© 2010 Mount Sinai School of Medicine

adults and children.[1] Hispanic adults and children had lower rates of lifetime asthma diagnosis, current asthma, and asthma attacks in comparison with non-Hispanic whites and non-Hispanic blacks, with an overall prevalence of 5.4%.[3] However, it is now clear that the Hispanic population in the United States is heterogeneous, and better characterization of this population will help improve our understanding of asthma in this population. In the US Census data, the term *Hispanic* refers to peoples with roots in Spain or in areas that were previously under Spanish control, including Mexico, much of Central and South America, and some Caribbean islands.[4] As shown in data from the 2000 US Census, the 35.2 million Hispanics in the US population derive from a wide range of countries and geographic areas: 59.3% of Hispanics in the United States were from Mexico; 22.3% traced their ancestry to Puerto Rico, Central America, South America, or Cuba; and 18.3% came from other areas.[5] Importantly, 47.9% of Hispanics characterized their race as white, and 42.2% characterized their race as other. The country of origin is relevant for asthma prevalence because it varies throughout Latin America, with Mexico having the lowest childhood prevalence and Costa Rica having the highest.[6]

Hispanic adults and children had lower rates of lifetime asthma diagnosis, current asthma, and asthma attacks in comparison with non-Hispanic whites and non-Hispanic blacks, with an overall prevalence of 5.4%.

This heterogeneity in asthma prevalence, symptoms, and severity in the Hispanic population has been seen in many studies, and indeed, on closer analysis of the 2000 US Census data, although Hispanics as a whole had a low asthma prevalence, there was great variation in the rates, with a prevalence of 5.0% in Mexican Americans and a prevalence of 17.0% in Puerto Ricans.[1]

The country of origin is relevant for asthma prevalence because it varies throughout Latin America, with Mexico having the lowest childhood prevalence and Costa Rica having the highest.

In the United States, 90% of the US Mexican population lives in the Southwest, and 68% of the US Puerto Rican population lives in the Northeast; the question could be raised whether the differences in asthma prevalence among Hispanics of varied ancestry are due to environmental exposures and confounders.[7] Studies within urban areas, however, suggest that environmental exposure incompletely explains this variation. In New York City, asthma prevalence is highest among Puerto Rican children,[8,9] and this increase in prevalence persists even after adjustments for the neighborhood and income level.[10] Puerto Rican adults in the New York City Community Health Survey have a higher asthma prevalence than other Hispanic subgroups and Hispanics born outside the United States, and the increased prevalence is only partially explained by age, sex, socioeconomic status, or other potential confounders.[11] In Chicago, the overall prevalence of asthma in children is 12.9%; however, there is variation between neighborhoods, and some but not all of the variation is explained by race.[12] An analysis of the National Health Interview Survey data for children in all 50 states of the United States shows that Puerto Rican children have the highest burden of asthma, whereas Mexican children have the lowest.[13] In contrast to other Hispanic subgroups in which birth outside the United States is associated with a lower risk for asthma, the prevalence for asthma in the Puerto Rican population is increased in island-born children.[13] The high prevalence of asthma in specific Hispanic subgroups reinforces the need to study asthma in these groups. Moreover, the heterogeneity of asthma prevalence in Hispanic subgroups suggests that epidemiological and genetic studies need to incorporate ancestral information as well as environmental exposures.

GENETIC STUDIES AND ASTHMA DISPARITIES

The differences in asthma prevalence with the country of origin suggest a complex relationship between ancestry and asthma susceptibility in the Hispanic population.

The differences in asthma prevalence with the country of origin suggest a complex relationship between ancestry and asthma susceptibility in the Hispanic population.

The association between Puerto Rican ethnicity and asthma prevalence and morbidity suggests that multiple factors, including genetic ones, may participate in the association. How these additional factors, including genetics, can promote the reduction in health disparities has been outlined in an excellent review.[14] As discussed in the review, rare variants, or susceptibility alleles with low frequency (≤2%) in the general population, may be found only in certain ethnic groups, and this reinforces the need to study diverse populations. Identification of these variants can provide specific insight into a disease in the selected ethnic population and may also provide insight into the biology of the disease. Second, disease susceptibility alleles that are common in some ethnic groups may be more or less common in other groups. Ample examples of this exist for genetic diseases such as cystic fibrosis, Tay-Sachs disease, and factor V Leiden, and screening for the presence of disease may be helped by an understanding of the allelic frequencies of variants in different populations. Third, genetic variants may affect therapy, and an understanding of the interactions between these genetic variants and specific therapeutic interventions as well as their prevalence in a population may have significant treatment implications. Finally, disease risk may be altered by the interaction between genetic variants and environmental interactions. Understanding this may be critical for disease prevention and may influence legislation for disease reduction. The Clean Air Act is an important example in which this potential exists.[15] Under the Clean Air Act, the Environmental Protection Agency sets National Ambient Air Quality Standards for 6 pollutants (carbon monoxide, lead, nitrogen dioxide, ozone, sulfur oxides, and particulate matter). Standards are based on the need to protect public health and the protection of sensitive human subpopulations. The definition of *sensitive human subpopulation* has been argued to include those identified by genomics (Interim Policy on Genomics, US Environmental Protection Agency. 2002). Thus, there is a potential role for genetic information to be included in the process for developing air quality standards.

COMPLEXITY OF STUDYING GENETICS, RACE AND ETHNICITY, AND HEALTH DISPARITIES IN THE LATINO POPULATION

The study of genetics in the Latino population is not a trivial matter. The expression of genetic variants is modified by huge environmental effects of pollution, allergen exposure, social stressors, and comorbid conditions. The relative importance and use of race/ethnicity in biomedical research and genetics are controversial, and the importance is debated.[16]

> *The expression of genetic variants is modified by huge environmental effects of pollution, allergen exposure, social stressors, and comorbid conditions. The relative importance and use of race/ethnicity in biomedical research and genetics are controversial, and the importance is debated.*

Moreover, there is not complete agreement on the definitions of *ethnicity* and *race*.[17] The concepts of race and ethnicity include connotations that reflect culture, history, socioeconomics, linguistics, and political status, although race has been considered to be characterized by the primary continent of origin.[17–20] Recent studies that assess the relationship of self-reported race and genetic ancestry with allergic sensitization in a non-Latino population provide conflicting data about the importance of the use of race/ethnicity; in one study, self-reported African American race was associated with allergic sensitization, asthma, and high total serum immunoglobulin E (IgE) levels,[21] whereas in another study, the association with allergic sensitization was not significant after researchers accounted for the location of residence.[22]

The study of genetics is even more complex and difficult when the study population is admixed, that is, contains race and ethnic mixtures.

> *The study of genetics is even more complex and difficult when the study population is admixed, that is, contains race and ethnic mixtures.*

The Hispanic community is an admixed population, and the allelic frequency differences among different race/ethnic groups can interfere with association studies and lead to false-positive results. The current Hispanic population is a result of combinations of individuals from Europe, Africa, and the Americas that resulted from intermating of migrant and indigenous populations.[5] This intermingling resulted in admixed populations and a combination of genomes that evolved independently on different continents.[23] Self-report of race and ancestry has been used extensively to assess race and ethnicity; however, self-report may be a poor proxy for underlying genetic relatedness,[16,24–26] and as recently described for an African American population, self-report may not predict ancestral clusters or reveal the extent of admixture.[27] In general, Puerto Ricans and Cubans have a greater proportion of African ancestry but a lower proportion of Native American ancestry than Mexican Americans.

The need to compensate for population stratification in genetic studies can be reduced if studies use family-based controls rather than case-control methods. Family-based control studies are often analyzed with transmission disequilibrium tests; however, this approach may not always be feasible because it requires first-degree relatives and usually genotyping of the proband and both parents. It is often difficult to recruit families, and in studies of adults or diseases with adult onset, it may be difficult to obtain parents. Thus, case-control studies are commonly used for genome-wide association (GWA) studies, candidate gene studies, and replication studies of GWA findings because of their cost effectiveness. However, in these studies, the possibility exists that allele frequency differences between cases and controls are unrelated to the genetics of the disease phenotype and instead are due to differences in the population genetic structure and substructure. Thus, false associations of particular single-nucleotide polymorphisms (SNPs) or markers can be identified, and each SNP may be affected by ancestry to a different extent, with some SNPS showing very different allele frequencies in different continental or ethnic groups and others showing little difference.

An approach to overcoming the problem of population substructure in genetic analyses makes use of DNA-based estimates of ancestry and appropriate statistical techniques that can add information to stratify populations and provide the ability to control for genetic differences that are unrelated to disease risk.[28–31] Ancestry informative markers (AIMs) are sets of polymorphisms for a locus that exhibit different frequencies between populations from different geographic regions. Preferably, AIMs are widely spaced across the genome so that there is no linkage disequilibrium among them within each subpopulation. Admixture mapping or mapping by admixture linkage disequilibrium takes advantage of sets of markers that are relatively evenly spaced across the genome. Several approaches, including structured association tests (Structure)[32] and principal component analyses,[33,34] have been proposed to adjust population stratifications in genetic association studies. Structure performs tests that are conditional on the group membership, which is determined with a Bayesian clustering algorithm that fits the data to the number of cluster groups that is specified. The association test is for the null hypothesis that there is no dependence of allele frequencies on phenotypes within each group. Principal component analysis infers a continuous axis of orthogonal linear transformations of AIMs that maximize the genetic variation but does not depend on the assignment of individuals to specific subpopulations. The number of SNPs is reduced to smaller numbers of dimensions that group patterns together on the basis of observed data. Ancestry-adjusted genotypes and phenotypes can then be calculated with Eigenstrat to compute the association statistic.[34] This algorithm can be difficult to perform with large numbers of SNPS and large populations. Plink, another commonly used program, can be used for large data sets[35] through the use of identical-by-state distance to do hierarchical clustering and subsequent Cochran-Mantel-Haenszel tests of association.

The use of AIMs has recently been applied to Hispanic populations, and variations in the ancestral composition of European Americans, Mexican Americans, disparate populations within the Native American population, Puerto Ricans, African Americans, and Southeast Asians have been demonstrated.[29,31,36–39] Studies using AIMs suggest that population admixture among Latinos has important implications when disease association is being assessed in case-control studies.[37] Analyses between these populations suggest that there are differences in the severity of asthma among Mexican Americans and Puerto Ricans and that there is an association between greater European ancestry in these populations and more severe asthma.[39]

GENETICS OF ASTHMA

Asthma is a complex genetic disorder that does not follow classic Mendelian inheritance; instead, each genetic factor contributes at most a modest amount to the total variation in the trait. Moreover, trait definition and phenotypic variation complicate the study of asthma. With the completion of the human genome sequence in 2005 and the development of techniques to assess thousands of SNPs, multiple techniques, including GWA studies, have been used to identify hundreds of genetic loci in which common genetic

Asthma is a complex genetic disorder that does not follow classic Mendelian inheritance; instead, each genetic factor contributes at most a modest amount to the total variation in the trait.

variations occur that are reproducibly associated with polygenic traits.[40] However, the analysis of these variations is complex, and the use of these variations to understand disease and disease biology is in its early stages for asthma as well as other diseases.[40,41] Extensive linkage analyses (positional cloning) and GWA studies, in which the entire genome is interrogated, as well as candidate gene association studies, in which a priori hypotheses are tested, have been performed for asthma, and several excellent reviews discussing both the successes and problems of the different approaches have recently been published.[42–45]

Studies using linkage analysis, a hypothesis-free approach to genetic associations, have identified 6 genes: a disintegrin and metalloproteinase 33 (*ADAM33*; chromosome 20p13), a metalloprotease; dipeptidyl peptidase 10 (chromosome 2q14), a regulatory molecule involved in protein targeting; plant homeodomain finger protein 11 (chromosome 13q14), a gene involved in messenger RNA transcription; G protein–coupled receptor for asthma (chromosome 7p14); human leukocyte antigen G (chromosome 6p21), a major histocompatibility complex antigen; and *GYFIP2* (chromosome 5q33), a G protein modulator.[42] Additional positional cloning studies have recently identified associations with genes predominantly expressed on epithelial cells: interleukin 1 receptor–associated kinase M (chromosome 12q14), a gene that encodes a protein regulating nuclear factor kappa B and the Toll-like receptor/interleukin 1 receptor pathway[46]; and opsin 3 (chromosome 1q43), a gene whose function remains unknown.[47] The role and function of many of these genes remain to be clarified.

GWA studies, in which large representative (tagged) marker panels of SNPs (100,000–1,000,000) are analyzed across the genome, are becoming more common as the technical ability and the density of markers available for study increase. The orosomucoid 1-like 3 (*ORMDL3*) gene (chromosome 17q21) was the first gene to be identified with childhood asthma through a GWA study.[48] The collaborative study also allowed for replication in additional British and German populations. Subsequent replication in a French-Canadian Caucasian population was also performed.[49] The gene that encodes YKL-40 (chitinase 3-like 1), a chitinase-like protein, has also been identified in a GWA study as a susceptibility gene for asthma, measures of lung function, and serum levels of YKL-40.[50]

Genetic association studies using case-control techniques for biologically plausible genes have identified numerous candidate SNPs and genes. The hundreds of genes identified with this approach have recently been reviewed.[44,45,51] Relatively few candidate gene associations have been replicated in multiple independent studies, and those that have been replicated repeatedly include the β-2 adrenergic receptor (*ADRB2*) gene, which is located in the cytokine gene cluster on chromosome 5q31-33, the interleukin 4 receptor, and the human leukocyte antigen DRB1 gene (chromosome 6).

Genetic studies of asthma are difficult because of the modest size of the genetic effects, which necessitates large sample sizes. Linkage analyses and GWA studies require sample sizes of thousands of individuals for successful identification of small effects and the need to compensate for multiple testing to enhance the ability to distinguish true positives from false-positive errors. In addition, for all of these genetic studies, replication of associated loci or candidate genes in independent populations is critical; however, replication has been variable between independent studies and in diverse populations for asthma genes.[52] This variability in replication has been attributed in part to differences in phenotypes and definitions of asthma-associated phenotypes, coverage of common genetic variants by SNPs, and population differences.

SELECTED ASTHMA GENES IN THE LATINO POPULATION

A limited number of genes associated with asthma risk have been identified in diverse Hispanic populations (Table 1). Regions on chromosome 5q, including regions 5q23 and 5q31-33, are some of the first well-replicated regions linked to asthma or asthma-associated phenotypes, and the association with this region has been reported in multiple candidate gene association studies in numerous populations, including Dutch, Amish, American Caucasian, Hutterite, and British.[44,45] The importance of this region or candidate genes for asthma in this region in Hispanic populations is less well documented. A recent GWA and admixture mapping

Table 1. *Asthma-Related Genes Studied in Hispanic Populations.*

Gene	Gene Map Locus	Biological Process	Hypothesized Role	References
IL-13	5q31	Th2 effector	Asthma risk, eosinophils, IgE	53–56
ADRB2	5q32-34	Bronchial smooth muscle function	Airway remodeling, lung function, asthma severity	57
ADAM33	20p13	Proteolysis, cell adhesion–mediated signaling	Asthma risk, lung function, airway remodeling	61–64
ORMDL3	17q21.1	Protein folding, calcium signaling	Childhood asthma risk	69, 71
TSLP	5q23	Epithelial cell Th2 regulation	Gender-specific IgE regulation	76

Abbreviations: ADAM33, a disintegrin and metalloproteinase 33; ADRB2, β-2 adrenergic receptor; IgE, immunoglobulin E; IL-13, interleukin 13; ORMDL3, orosomucoid 1-like 3; Th2, T helper 2; TSLP, thymic stromal lymphopoietin.

analysis for asthma in a very small Puerto Rican sample failed to identify any association with asthma in this region, although a subsequent overlap analysis suggested an association in 5q23.3 as well as 13q.13.3.[53] Many genes in this 5q area, including the gene for interleukin 13 (IL-13), are in the cytokine cluster associated with Th2 cell differentiation and effector functions. IL-13 is a central effector cytokine and mediates many of the features of Th2 cell–mediated responses in the lung, including IgE synthesis. A study of a Costa Rican population initially failed to identify an association with IL-13 and asthma[54]; however, a subsequent study of Costa Rican children, using a larger sample size and a replication group, suggested that some IL-13 SNPS were associated with an increased eosinophil count, increased total serum IgE levels, and asthma exacerbation.[55] Our preliminary data, coming from a case-control association study for IL-13 that controlled for ancestral diversity through the use of AIMs in a predominantly Latino, admixed urban population, also identified an association with the frequently reported IL-13 variant.[56] These data suggest that findings in IL-13 variants may be robust even within different Latino populations, but additional studies with large populations are needed.

The *ADRB2* gene has been classified as a gene involved in lung function, airway remodeling, and disease severity.[44] Variants in 7 SNPS in the *ADRB2* gene were analyzed in the Genetics of Asthma in Latino Americans study with Puerto Rican and Mexican families of asthmatics[57] Although 2 SNPS were associated with the baseline forced expiratory volume in 1 second and bronchodilator responsiveness among the Puerto Rican families, a different SNP was marginally associated with bronchodilator responsiveness in the Mexican families. These studies again suggest differences between Latino populations and the need for additional studies.

ADAM33, a disintegrin and metalloproteinase gene located on chromosome 20, was the first asthma susceptibility gene to be identified through linkage studies and positional cloning.[58] Subsequent studies suggested a potential role for this gene in airway remodeling and revealed that *ADAM33* is expressed by lung fibroblasts and bronchial smooth muscle cells, although not by bronchial epithelial cells or immune cells.[59] *ADAM33* variants have been associated with an accelerated decline in the forced expiratory volume in 1 second.[60] Although the association of *ADAM33* variants and asthma has been replicated in many different ethnic populations,[43] the association of *ADAM33* and asthma in Latino populations has been less clear. One study using diverse populations, which included 112 Hispanic cases, suggested that variations in *ADAM33* were associated with asthma in all populations, although differences were noted between populations, and no causative variation could be identified in all populations.[61] A weak association with *ADAM33* variants and asthma was noted in children in the Childhood Asthma Management Program, a study that included only 39 children of Hispanic ethnicity.[62] An association between asthma and *ADAM33* variants was not found in Puerto Rican or Mexican adults,[63] nor was an association identified in Costa Rican children.[64] These studies differ in the techniques used for analysis, the size of the population studied, and the ancestry of the Latino populations. However, they suggest that the association of *ADAM33* with asthma remains inconsistent in the Latino population.

The *ORMDL3* gene was the first gene to be identified with a GWA study in white British and German patients.[48] *ORMDL3* transcription levels are correlated with several genetic markers near the *ORMDL3* gene, and although the function of *ORMDL3* remains incompletely understood, the gene product may participate in protein folding, calcium signaling, and cellular stress.[65] Polymorphisms in the *ORMDL3* gene have been associated with childhood asthma in multiple ethnic populations.[66–68] Two studies suggest an association with Latino populations. A recent study reported an association of *ORMDL3* gene variants and asthma in Mexican children.[69] In another, an association was identified in a Mexican population, but only a weak association was found in Puerto Rican children.[70] The possibility that smoke exposure modifies the risk for *ORMDL3* variants and asthma raises an intriguing question about environmental interactions.[71]

Our interest has focused on the ability of airway epithelial cells to regulate airway immune responses, particularly in response to environmental injury by pollutants. Thymic stromal lymphopoietin (TSLP) is an epithelial cell–derived cytokine that regulates the Th2 response,[72–74] and we have described its production in response to exposure to diesel exhaust particles.[75] Thus, the gene for TSLP, located on chromosome 5, is a biologically plausible candidate gene for the study of gene and environment interactions. A locus on 5q23 was associated with cockroach-specific IgE in Costa Rican children in a genome-wide linkage analysis when it was analyzed with sex-stratified linkage analysis.[76] The TSLP gene was identified in this region, and the proposal was made that a variant in TSLP may have female-specific effects on allergic phenotypes. Further replication studies are needed to confirm this association in additional populations, but the biological importance of TSLP makes the study provocative.

CONCLUSION

In summary, the high prevalence of asthma in specific Latino subpopulations makes it imperative to study asthma in Hispanics. The study of asthma and asthma genetics is made complex in these populations by the diversity and heterogeneity of the Hispanic population in the United States. The inclusion of markers of ancestral diversity and the incorporation of techniques to adjust for stratification now make these studies more feasible and will enable an understanding of the underlying biology of asthma, genetic risks for asthma, and gene and environmental interactions. These studies have the potential to enhance risk identification, target pharmacological therapy, guide environmental regulations, and thus reduce health disparities. New technical advances that now allow for the study of asthma and asthma-associated phenotypes at a rapid pace using linkage analyses and GWA studies as well as candidate gene studies require large populations. Although large populations of European descent exist for genetic studies of asthma, fewer well-phenotyped populations are available for studies of Latino populations. There is a need for the development of registries with large pediatric and adult Latino populations for inclusion in the rapidly expanding field of genetic and environmental studies. These studies will promote the ultimate goal of improving our understanding of the biology and treatment of asthma in all populations and thus reduce health disparities.

ACKNOWLEDGMENT

This study was supported by the Colton Family Foundation, the National Institutes of Health/National Center for Research Resources (grant M01RR00096), the Stony Wold-Herbert Fund, Inc., and the National Institute of Environmental Health Sciences (grant 000260). The material was presented in part at a conference on genes and race held at the Mt. Sinai Medical Center in November 2008.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Centers for Disease Control and Prevention. FastStats on asthma. http://www.cdc.gov/nchs/FASTATS/asthma.htm. Accessed January 2010.
2. Garg R, Karpati A, Leighton J, et al. *Asthma Facts*. 2nd ed. New York, NY: New York City Department of Health and Mental Hygiene; 2003.
3. Moorman JE, Rudd RA, Johnson CA, et al. National surveillance for asthma–United States, 1980–2004. *MMWR Surveill Summ* 2007; 56: 1–54.
4. *We the People: Hispanics in the United States: Census 2000 Special Report*. Washington, DC: US Census Bureau; 2004.
5. Hunninghake GM, Weiss ST, Celedon JC. Asthma in Hispanics. *Am J Respir Crit Care Med* 2006; 173: 143–163.
6. Pearce N, Ait-Khaled N, Beasley R, et al. Worldwide trends in the prevalence of asthma symptoms: phase III of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax* 2007; 62: 758–766.
7. US Census Bureau. Hispanic population in the United States: March 2002. http://www.census.gov/prod/2003pubs/p20-545.pdf. Accessed January 2010.
8. Cloutier MM, Wakefield DB, Hall CB, Bailit HL. Childhood asthma in an urban community: prevalence, care system, and treatment. *Chest* 2002; 122: 1571–1579.
9. Findley S, Lawler K, Bindra M, et al. Elevated asthma and indoor environmental exposures among Puerto Rican children of East Harlem. *J Asthma* 2003; 40: 557–569.
10. Claudio L, Stingone JA, Godbold J. Prevalence of childhood asthma in urban communities: the impact of ethnicity and income. *Ann Epidemiol* 2006; 16: 332–340.

11. Dumanovsky T, Matte TD. Variation in adult asthma prevalence in Hispanic subpopulations in New York City. *J Asthma* 2007; 44: 297–303.
12. Subramanian SV, Jun HJ, Kawachi I, Wright RJ. Contribution of race/ethnicity and country of origin to variations in lifetime reported asthma: evidence for a nativity advantage. *Am J Public Health* 2009; 99: 690–697.
13. Lara M, Akinbami L, Flores G, Morgenstern H. Heterogeneity of childhood asthma among Hispanic children: Puerto Rican children bear a disproportionate burden. *Pediatrics* 2006; 117: 43–53.
14. Scirica CV, Celedon JC. Genetics of asthma: potential implications for reducing asthma disparities. *Chest* 2007; 132(suppl): 770S–781S.
15. Kramer CB, Cullen AC, Faustman EM. Policy implications of genetic information on regulation under the Clean Air Act: the case of particulate matter and asthmatics. *Environ Health Perspect* 2006; 114: 313–319.
16. Caulfield T, Fullerton SM, Ali-Khan SE, et al. Race and ancestry in biomedical research: exploring the challenges. *Genome Med* 2009; 1: 8.
17. Collins FS. What we do and don't know about "race", "ethnicity", genetics and health at the dawn of the genome era. *Nat Genet* 2004; 36(suppl): S13–S15.
18. Bamshad M, Wooding S, Salisbury BA, Stephens JC. Deconstructing the relationship between genetics and race. *Nat Rev Genet* 2004; 5: 598–609.
19. Risch N, Burchard E, Ziv E, Tang H. Categorization of humans in biomedical research: genes, race and disease. *Genome Biol* 2002; 3: 1–12.
20. Cooper RS, Kaufman JS, Ward R. Race and genomics. *N Engl J Med* 2003; 348: 1166–1170.
21. Vergara C, Caraballo L, Mercado D, et al. African ancestry is associated with risk of asthma and high total serum IgE in a population from the Caribbean coast of Colombia. *Hum Genet* 2009; 125: 565–579.
22. Yang JJ, Burchard EG, Choudhry S, et al. Differences in allergic sensitization by self-reported race and genetic ancestry. *J Allergy Clin Immunol* 2008; 122: 820–827.
23. Basu A, Tang H, Zhu X, et al. Genome-wide distribution of ancestry in Mexican Americans. *Hum Genet* 2008; 124: 207–214.
24. Suarez-Kurtz G, Vargens DD, Struchiner CJ, et al. Self-reported skin color, genomic ancestry and the distribution of GST polymorphisms. *Pharmacogenet Genomics* 2007; 17: 765–771.
25. Serre D, Paabo S. Evidence for gradients of human genetic diversity within and among continents. *Genome Res* 2004; 14: 1679–1685.
26. Wilson JF, Weale ME, Smith AC, et al. Population genetic structure of variable drug response. *Nat Genet* 2001; 29: 265–269.
27. Yaeger R, Avila-Bront A, Abdul K, et al. Comparing genetic ancestry and self-described race in African Americans born in the United States and in Africa. *Cancer Epidemiol Biomarkers Prev* 2008; 17: 1329–1338.
28. Hoggart CJ, Parra EJ, Shriver MD, et al. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet* 2003; 72: 1492–1504.
29. Yang N, Li H, Criswell LA, et al. Examination of ancestry and ethnic affiliation using highly informative diallelic DNA markers: application to diverse and admixed populations and implications for clinical epidemiology and forensic medicine. *Hum Genet* 2005; 118: 382–392.
30. Tsai HJ, Kho JY, Shaikh N, et al. Admixture-matched case-control study: a practical approach for genetic association studies in admixed populations. *Hum Genet* 2006; 118: 626–639.
31. Collins-Schramm HE, Phillips CM, Operario DJ, et al. Ethnic-difference markers for use in mapping by admixture linkage disequilibrium. *Am J Hum Genet* 2002; 70: 737–750.
32. Satten GA, Flanders WD, Yang Q. Accounting for unmeasured population substructure in case-control studies of genetic association using a novel latent-class model. *Am J Hum Genet* 2001; 68: 466–477.
33. Tian C, Gregersen PK, Seldin MF. Accounting for ancestry: population substructure and genome-wide association studies. *Hum Mol Genet* 2008; 17():R143R150.
34. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; 38: 904–909.
35. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559–575.
36. Collins-Schramm HE, Chima B, Morii T, et al. Mexican American ancestry-informative markers: examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians and Asians. *Hum Genet* 2004; 114: 263–271.
37. Choudhry S, Coyle NE, Tang H, et al. Population stratification confounds genetic association studies among Latinos. *Hum Genet* 2006; 118: 652–664.
38. Mao X, Bigham AW, Mei R, et al. A genomewide admixture mapping panel for Hispanic/Latino populations. *Am J Hum Genet* 2007; 80: 1171–1178.
39. Salari K, Choudhry S, Tang H, et al. Genetic admixture and asthma-related phenotypes in Mexican American and Puerto Rican asthmatics. *Genet Epidemiol* 2005; 29: 76–86.
40. Hirschhorn JN. Genomewide association studies–illuminating biologic pathways. *N Engl J Med* 2009; 360: 1699–1701.
41. Goldstein DB. Common genetic variation and human traits. *N Engl J Med* 2009; 360: 1696–1698.
42. Willis-Owen SA, Cookson WO, Moffatt MF. Genome-wide association studies in the genetics of asthma. *Curr Allergy Asthma Rep* 2009; 9: 3–9.
43. Weiss ST, Raby BA, Rogers A. Asthma genetics and genomics 2009. *Curr Opin Genet Dev* 2009; 19: 279–282.
44. Vercelli, D. Advances in asthma and allergy genetics in 2007. *J Allergy Clin Immunol* 2008; 122: 267–271.
45. Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immunol* 2006; 7: 95–100.
46. Balaci L, Spada MC, Olla N, et al. IRAK-M is involved in the pathogenesis of early-onset persistent asthma. *Am J Hum Genet* 2007; 80: 1103–1114.
47. White JH, Chiano M, Wigglesworth M, et al. Identification of a novel asthma susceptibility gene on chromosome 1qter and its functional evaluation. *Hum Mol Genet* 2008; 17: 1890–1903.

48. Moffatt MF, Kabesch M, Liang L, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007; 448: 470–473.
49. Madore AM, Tremblay K, Hudson TJ, Laprise C. Replication of an association between 17q21 SNPs and asthma in a French-Canadian familial collection. *Hum Genet* 2008; 123: 93–95.
50. Ober C, Tan Z, Sun Y, et al. Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 2008; 358: 1682–1691.
51. Zhang J, Pare PD, Sandford AJ. Recent advances in asthma genetics. *Respir Res* 2008; 9: 4.
52. Lohmueller KE, Pearce CL, Pike M, et al. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003; 33: 177–182.
53. Choudhry S, Taub M, Mei R, et al. Genome-wide screen for asthma in Puerto Ricans: evidence for association with 5q23 region. *Hum Genet* 2008; 123: 455–468.
54. Celedon JC, Soto-Quiros ME, Palmer LJ, et al. Lack of association between a polymorphism in the interleukin-13 gene and total serum immunoglobulin E level among nuclear families in Costa Rica. *Clin Exp Allergy* 2002; 32: 387–390.
55. Hunninghake GM, Soto-Quiros ME, Avila L, et al. Polymorphisms in IL13, total IgE, eosinophilia, and asthma exacerbations in childhood. *J Allergy Clin Immunol* 2007; 120: 84–90.
56. Cheng Q, Liu M, Rogers L, et al. Replication study of association of IL13 pathway polymorphisms and asthma using an admixed urban population. *Am J Respir Crit Care Med* 2009; 179.
57. Choudhry S, Ung N, Avila PC, et al. Pharmacogenetic differences in response to albuterol between Puerto Ricans and Mexicans with asthma. *Am J Respir Crit Care Med* 2005; 171: 563–570.
58. Van Eerdewegh P, Little RD, Dupuis J, et al. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 2002; 418: 426–430.
59. Haitchi HM, Powell RM, Shaw TJ, et al. ADAM33 expression in asthmatic airways and human embryonic lungs. *Am J Respir Crit Care Med* 2005; 171: 958–965.
60. Jongepier H, Boezen HM, Dijkstra A, et al. Polymorphisms of the ADAM33 gene are associated with accelerated lung function decline in asthma. *Clin Exp Allergy* 2004; 34: 757–760.
61. Howard TD, Postma DS, Jongepier H, et al. Association of a disintegrin and metalloprotease 33 (ADAM33) gene with asthma in ethnically diverse populations. *J Allergy Clin Immunol* 2003; 112: 717–722.
62. Raby BA, Silverman EK, Kwiatkowski DJ, et al. ADAM33 polymorphisms and phenotype associations in childhood asthma. *J Allergy Clin Immunol* 2004; 113: 1071–1078.
63. Lind DL, Choudhry S, Ung N, et al. ADAM33 is not associated with asthma in Puerto Rican or Mexican populations. *Am J Respir Crit Care Med* 2003; 168: 1312–1316.
64. Hersh CP, Raby BA, Soto-Quiros ME, et al. Comprehensive testing of positionally cloned asthma genes in two populations. *Am J Respir Crit Care Med* 2007; 176: 849–857.
65. Cantero-Recasens G, Fandos C, Rubio-Moscardo F, et al. The asthma-associated ORMDL3 gene product regulates endoplasmic reticulum-mediated calcium signaling and cellular stress. *Hum Mol Genet* 2010; 19: 111–121.
66. Sleiman PM, Annaiah K, Imielinski M, et al. ORMDL3 variants associated with asthma susceptibility in North Americans of European ancestry. *J Allergy Clin Immunol* 2008; 122: 1225–1227.
67. Hirota T, Harada M, Sakashita M, et al. Genetic polymorphism regulating ORM1-like 3 (*Saccharomyces cerevisiae*) expression is associated with childhood atopic asthma in a Japanese population. *J Allergy Clin Immunol* 2008; 121: 769–770.
68. Leung TF, Sy HY, Ng MC, et al. Asthma and atopy are associated with chromosome 17q21 markers in Chinese children. *Allergy* 2009; 64: 621–628.
69. Wu H, Romieu I, Sienra-Monge JJ, et al. Genetic variation in ORM1-like 3 (ORMDL3) and gasdermin-like (GSDML) and childhood asthma. *Allergy* 2009; 64: 629–635.
70. Galanter J, Choudhry S, Eng C, et al. ORMDL3 gene is associated with asthma in three ethnically diverse populations. *Am J Respir Crit Care Med* 2008; 177: 1194–1200.
71. Flory JH, Sleiman PM, Christie JD, et al. 17q12-21 variants interact with smoke exposure as a risk factor for pediatric asthma but are equally associated with early-onset versus late-onset asthma in North Americans of European ancestry. *J Allergy Clin Immunol* 2009; 124: 605–607.
72. Rochman Y, Leonard WJ. Thymic stromal lymphopoietin: a new cytokine in asthma. *Curr Opin Pharmacol* 2008; 8: 249–254.
73. Liu YJ. Thymic stromal lymphopoietin: master switch for allergic inflammation. *J Exp Med* 2006; 203: 269–273.
74. Soumelis V, Liu YJ. Human thymic stromal lymphopoietin: a novel epithelial cell-derived cytokine and a potential key player in the induction of allergic inflammation. *Springer Semin Immunopathol* 2004; 25: 325–333.
75. Bleck B, Tse DB, Curotto de Lafaille MA, et al. Diesel exhaust particle-exposed human bronchial epithelial cells induce dendritic cell maturation and polarization via thymic stromal lymphopoietin. *J Clin Immunol* 2008; 28: 147–156.
76. Hunninghake GM, Lasky-Su J, Soto-Quiros ME, et al. Sex-stratified linkage analysis identifies a female-specific locus for IgE to cockroach in Costa Ricans. *Am J Respir Crit Care Med* 2008; 177: 830–836.

Is There a Genetic Basis for Health Disparities in Human Immunodeficiency Virus Disease?

Cheryl Winkler, PhD

Laboratory of Genomic Diversity, National Cancer Institute at Frederick, Frederick, MD

OUTLINE

Chemokine (C-C Motif) Receptor 5 Δ32, Human Immunodeficiency Virus Acquisition, and Progression to Acquired Immune Deficiency Syndrome
Duffy Antigen Chemokines and HIV Acquisition
Hiv-Asscoaited Nephropathy AND *MYH9*
Genes That Modify Human Immunodeficiency Virus
Conclusion

ABSTRACT

The highest global prevalence rates for human immunodeficiency virus and acquired immune deficiency syndrome have been recorded in southern Africa; in the United States, individuals of African descent are disproportionately affected by human immunodeficiency virus infection. Human immunodeficiency virus–infected individuals with African ancestry are also estimated to have a 17-fold or greater risk for developing human immunodeficiency virus–associated nephropathy in comparison with their counterparts of non-African descent. Several recent studies have implicated genetic alleles that are more frequent in populations of African descent and increase the risk of human immunodeficiency virus infection and the risk of human immunodeficiency virus–associated neuropathy (HIVAN). The supposition that persons of African descent are more susceptible to human immunodeficiency virus infection because of an underlying genetic predisposition is not supported by available evidence. However, strong, replicated data show that the increased risk for human immunodeficiency virus–associated nephropathy, as well as other major forms of kidney disease in individuals of African descent, is due in part to *MYH9* (myosin, heavy chain 9, non-muscle) renal disease susceptibility alleles that are very frequent throughout sub-Saharan Africa but are infrequent or absent in non-Africans. Selection, drift, and demographic events shape the allelic architecture of the human genome: it is expected that these events will be reflected in geographic-specific differentiation in allele frequencies for a small subset of alleles that may be associated with either increased or reduced risk for complex and infectious diseases. ***Mt Sinai J Med 77:149–159, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: acquired immune deficiency syndrome (AIDS), health disparities, human immunodeficiency virus (HIV), human immunodeficiency virus–associated nephropathy (HIVAN), genetic alleles, genetic predisposition, myosin, heavy chain 9, non-muscle (*MYH9*), nephropathy, chromosome 22 renal susceptibility alleles.

Address Correspondence to:

Cheryl Winkler
Laboratory of Genomic Diversity
National Cancer Institute at Frederick
Frederick, MD
Email: winklerc@mail.nih.gov

Health disparities in the United States and worldwide are well documented for human immunodeficiency virus (HIV) infection, the development of acquired immune deficiency syndrome (AIDS), and AIDS-related deaths.[1–3] Although many of the inequities in HIV disease burden are due to nongenetic factors that include risk behaviors, access to health care, and socioeconomic status, host genetics may also play a role in geographic-specific or racial differences in

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20172

© 2010 Mount Sinai School of Medicine

disease susceptibility. Over the last decade, there has been an accumulation of data indicating that 12% to 15% of the variation in HIV progression and viral load is due to host genetic factors.[4,5] However, our knowledge of host genetic factors affecting AIDS is biased toward factors prevalent in whites. Because AIDS and HIV were first identified in North America and Europe and because of the early demographics of HIV infection, with the very first infections in these regions occurring in men who have sex with men and a short time later in men with hemophilia, most genetic association studies, including the more recent genome-wide association studies, have been focused on European or European American men, who are not at all representative of the current HIV global epidemic: more than half of HIV-infected persons live in sub-Saharan Africa, and of these, more than half are women.[1]

> *Over the last decade, there has been an accumulation of data indicating that 12% to 15% of the variation in HIV progression and viral load is due to host genetic factors.*

The Joint United Nations Program on HIV/AIDS reported that in 2007, of the 33 million people living with HIV infection, more than half (68%) were in sub-Saharan Africa, and 61% were women.[1] Eight countries in southern African account for 35% of HIV infections and 38% of AIDS-related deaths.[1] In the United States, HIV infection disproportionately affects blacks. Although blacks represent only 12% to 13% of the US population, nearly half (48%) of the 1.1 million persons with HIV infection share African ancestry. Blacks in the United States also account for a higher proportion of cases at all stages from infection to death with AIDS.[2] The small number of natural history and treatment cohort studies enrolling black or Hispanic subjects for genetic studies has limited the investigation of genetic predictors for HIV-1 infection and progression in these populations. Indeed, the major genome-wide association studies to date have enrolled European or European American participants.[5–9]

> *Eight countries in southern African account for 35% of HIV infections and 38% of AIDS-related deaths. In the United States, HIV infection disproportionately affects blacks. Although blacks represent only 12% to 13% of the US population, nearly half (48%) of the 1.1 million persons with HIV infection share African ancestry.*

Associations between HIV infection and a spectrum of renal glomerular diseases were first recognized in the mid 1980s. In 1984, Rao *et al.*[10] reported that HIV-infected patients were presenting with proteinuria in the nephrotic range followed by rapid progression to end-stage renal disease (ESRD), which was mainly attributable to focal and segmental glomerulosclerosis (FSGS). In European, Hispanic, and Asian patients, the most common form of HIV-associated glomerular disease is immune complex gomerulopnephritis,[11,12] whereas patients of African descent most frequently develop HIV-associated collapsing glomerulopathy, which is also known as human immunodeficiency virus–associated nephropathy (HIVAN). In 1995, HIVAN was the leading cause of ESRD in men of African descent between 20 and 64 years of age.[13] HIVAN is estimated to account for up to 30% of patients in the United States entering dialysis because of end-stage kidney disease as a result of FSGS.[14]

Notably, HIVAN is rarely diagnosed in people of European descent. The risk of developing HIVAN for individuals of African descent is at least 18-fold higher in comparison with the risk for their white counterparts; over 90% of persons with HIVAN are of African descent.[14] This predilection for FSGS has also been observed in HIV-infected people of African descent in Africa, Brazil, and France.[15–17] A recent review of the spectrum of kidney diseases in regions of Africa with high HIV prevalence rates also reports an unusually high prevalence of kidney disease.[18] Although no systematic study has been performed with standardized criteria for HIV-associated kidney disease, kidney disease among HIV-1–infected persons has been reported to range from 6% in South Africa to 38% in Nigeria.[18,19] Like HIVAN in developed nations, African ancestry is a risk predictor of HIVAN in sub-Saharan Africa.[18,19] Because of the clustering of kidney diseases with different etiologies (eg, HIVAN, hypertension-associated ESRD, and diabetes-associated ESRD) in African American families and the overall increased rates of many kidney diseases in African Americans, it has been hypothesized that renal disease susceptibility alleles may be more frequent in persons of African origin or descent.[20]

The dispersal of humans out of Africa 100,000 years ago was followed by a rapid expansion of human populations from a relatively small effective population size. This had the effect of limiting allele and haplotype diversity in non-African populations with respect to African populations: the extent of human genetic variation outside Africa is a subset of the genetic variation observed within Africa. There are very few fixed differences among the world's subpopulations: genetic diversity in the world's population is largely attributable to geographic differentials in allele frequencies.[21] The extent of genetic variation and diversity among continental, ethnic, and tribal groups reflects ancient, ongoing, and recent selective pressures from pathogens and environmental factors as well as historical and demographic events, such as migration and founder effects, on a background of random genetic drift.[22–24] The events shaping our human genome are complex, but they leave their footprints in the allelic and haplotype structure of our genomes: it should not be surprising that human continental populations, isolated subpopulations, and racial/ethnic groups may reflect this variation in differential genetic susceptibility to a spectrum of human pathogens and diseases.[25]

Here we consider the role of host genetic factors in population-specific differences in susceptibility to HIV infection and progression and in the development of HIVAN in African and European Americans. Because most AIDS-restricted genes have been extensively reviewed elsewhere, the focus of this commentary is a small subset of genes that are highly stratified between populations of European and African descent, that is, those with moderate to high allele frequencies (>10%) in one population but very low allele frequencies in the other; these may be expected to have large population-attributable risks that may provide a genetic basis for disparate risk between individuals of African descent and individuals of European descent. Rare polymorphisms with large effect sizes or common alleles with small effect sizes would not be expected to contribute substantially to HIV-related health disparities.

CHEMOKINE (C-C MOTIF) RECEPTOR 5 Δ32, HUMAN IMMUNODEFICIENCY VIRUS ACQUISITION, AND PROGRESSION TO ACQUIRED IMMUNE DEFICIENCY SYNDROME

The primary coreceptor used by transmissible strains of HIV to enter cells and establish infections is chemokine (C-C motif) receptor 5 (CCR5). In 1996, a series of articles reported that a 32–base pair deletion in the *CCR5* gene (*CCR5* Δ32) prevented HIV infection in cells from individuals homozygous for the mutated gene.[26–29] The mutation introduced a premature stop codon that truncated the CCR5 protein and effectively prevented its cell surface expression. The absence of CCR5 due to this deletion provided nearly complete resistance to HIV infection in those carrying 2 copies of the mutated gene.

> *The primary coreceptor used by transmissible strains of HIV to enter cells and establish infections is chemokine (C-C motif) receptor 5 (CCR5). A 32–base pair deletion in the CCR5 gene (CCR5 Δ32) prevented HIV infection in cells from individuals homozygous for the mutated gene. The absence of CCR5 due to this deletion provided nearly complete resistance in those carrying 2 copies of the mutated gene.*

CCR5 was the first HIV restriction gene discovered, and it remains to date the only mutation that protects against HIV acquisition. Although subsequent studies have identified a handful of individuals homozygous for *CCR5* Δ32 who have become HIV-infected, these infection events are extremely rare. Most HIV isolates from *CCR5* [delta]32-homozygous infected individuals have been found to use an alternative coreceptor, CXCR4, to establish infection.[30] These studies have established the importance of CCR5 as the primary coreceptor for transmissible strains of HIV.

Surveys of the population distribution of the *CCR5* Δ32 mutation have revealed that the deletion mutation is restricted to populations of northern European descent or European admixture.[31]

> *Surveys of the population distribution of the CCR5 Δ32 mutation have revealed that the deletion mutation is restricted to populations of northern European descent or European admixture.*

The frequencies of *CCR5 Δ32* are distributed across Europe in a north-to-south cline, with the highest frequencies in northern Europe (15% in Scandinavia, 10% in France, England, and Germany, and 5% in Italy, Greece, and Turkey), and it is rare to absent elsewhere in the world.[31,32] Estimates of the age of *CCR5 Δ32* suggest a relatively recent origin.[31] It has been speculated that the *CCR5 Δ32* mutation rose to its current frequency of approximately 10% in the European population by recent selection, possibly by a pathogen such as *Variola vaccinia* (smallpox) or *Yersina pestis* (black plague).[33]

Although homozygosity for *CCR5 Δ32* provides nearly absolute protection against HIV acquisition, this protection is afforded to only about 1% of northern Europeans homozygous for *CCR5 Δ32*, and this is too small a fraction to account for differences in infection rates among African and European individuals. Heterozygosity for *CCR5 Δ32* has been shown in many studies to delay the onset of AIDS by 2 or 3 years in untreated persons, and this protective effect continues after the initiation of antiretroviral therapy.[29,34–36] Carriers of *CCR5 Δ32* tend to have a lower HIV viral copy number; the HIV copy number of the transmitting partner or mother is positively correlated with HIV transmission.[36] It is conceivable that lower levels of HIV due to *CCR5 Δ32* or other polymorphisms in the *CCR5* gene that affect *CCR5* expression and protein levels may decrease HIV transmission efficiency and thereby affect the dynamics of HIV transmission in a population-specific manner.[37] Approximately 20% of Europeans are predicted to carry at least 1 *CCR5 Δ32* allele.

DUFFY ANTIGEN CHEMOKINES AND HIV ACQUISITION

The Duffy antigen receptor for chemokines (DARC) is a chemokine receptor primarily expressed on the surface of red blood cells. DARC has been extensively investigated because it is the cell surface receptor used by the malarial pathogen *Plasmodium vivax* to enter and infect red blood cells.[38] DARC provides classic evidence for the influence of human pathogens on natural selection.[39,40] Outside Africa, *P. vivax* is the most frequent cause of malaria in humans, but within Africa, malarial infections from *P. vivax* are extremely rare. A single base pair mutation in the *DARC* gene promoter (−46T→C) disrupts an erythrocyte transcription factor binding site required for *DARC* gene expression in erthrocytes.[41] As a result, individuals homozygous for the −46 C allele express no DARC receptor on their red blood cells and are resistant to *P. vivax* malaria. The *DARC* −46 C mutation is one of only a few mutations discovered that reaches near fixation (a >90% allele frequency in sub-Saharan Africa) in 1 continental population and is virtually absent in all other human populations. With an allele frequency of >90% in much of the malarial belt, most sub-Saharan Africans (≈81%) do not express DARC on their cell surfaces and are therefore resistant to *P. vivax* malaria.

DARC promiscuously absorbs a number of pro-inflammatory chemokines to the surfaces of red blood cells, including those that bind to the *CCR5* coreceptor and inhibit HIV binding to *CCR5* surface receptors.[42,43] The DARC cell surface receptor has also been reported to bind HIV to the surfaces of erythrocytes that may act to facilitate the transfer of HIV to susceptible cells.[44,45] Given the critical role of the chemokine/chemokine receptor nexus in HIV infection and the complete absence of the DARC receptor on red blood cells of most Africans, *DARC* is a highly plausible candidate gene for a role in modifying HIV acquisition and pathogenesis.

> *The Duffy antigen receptor for chemokines (DARC) is a chemokine receptor primarily expressed on the surface of red blood cells. Given the critical role of the chemokine/chemokine receptor nexus in HIV infection and the complete absence of the DARC receptor on red blood cells of most Africans, DARC is a highly plausible candidate gene for a role in modifying HIV acquisition and pathogenesis.*

He *et al.*[44] recently reported that individuals who do not express DARC on their cell surfaces are more likely to acquire HIV infection but survive longer after HIV infection. These observations were based on a

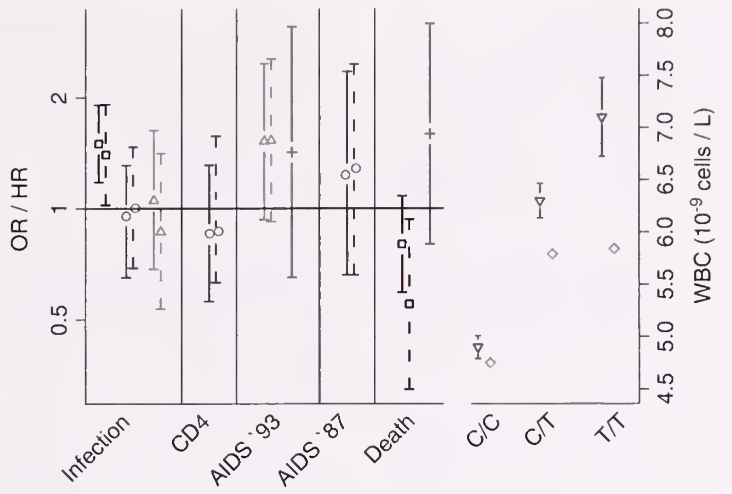


Fig 1. OR and HR values with 95% confidence intervals for associations of *DARC* with human immunodeficiency virus infection: time to CD4 <200; time to AIDS-defining condition (AIDS '93) or CD4 <200; and death from an AIDS-related illness. Crude and adjusted OR/HR values are represented by solid and dashed bars, respectively (left panel). Mean WBCs with 95% confidence intervals, where available, are also shown for each DARC genotype (right panel). The following studies are cited: (□) He *et al.*,[44] (○) Winkler *et al.*,[50] (△) Walley *et al.*,[49] (+) Horne *et al.*,[52] (◇) Julg *et al.*,[51] and (▽) Nalls.[53] **Abbreviations:** AIDS, acquired immune deficiency syndrome; DARC, Duffy antigen receptor for chemokines; HR, hazard ratio; OR, odds ratio; WBC, white blood cell count.

genetic association study enrolling African American men serving in the US military. The study was notable because the authors reported that the DARC null phenotype was associated with a 40% increase in the odds of acquiring HIV; this was based on the observation that the rate of the −46C/C genotype (abrogating DARC gene expression) was 70% in the HIV-positive group versus 60% in normal controls without HIV infection. The authors concluded that the DARC null genotype might account for 11% of the HIV-1 burden in Africa and increase the risk of HIV infection by 40%.[44] If correct, these results would suggest that individuals of African descent have a genetic vulnerability to HIV-1 acquisition; such a finding could have tremendous implications for public health policies aimed at stemming the tide of new HIV infections.

Such a finding warrants careful consideration and validation. African Americans are an admixed group with various degrees of African and European ancestry. It is estimated that 15% to 20% of the genes in African Americans in the United States derive from the European population.[46] Stratification between HIV-infected cases and HIV-uninfected controls must be carefully controlled to avoid a population substructure that may create spurious or inflated genetic associations between alleles and an outcome.[47,48] This is critical when a gene that is highly frequent in one population is rare or absent in a second population, such as the case with the *DARC* −46C null mutation. Four independent groups were unable to replicate the findings of He *et al.*[44] in individuals of African descent enrolled in United States–based natural history cohorts or in black Africans living in Durban, South Africa.[49–52] These results are summarized in Figure 1. These investigations varied somewhat in design and methods, but none of the studies provided support for He *et al.*'s finding that DARC was associated with either susceptibility to HIV infection or with the rate of progression to AIDS. For example, in Winkler *et al.*'s study of African American injecting drug users enrolled in a natural history cohort,[50] the frequencies of the −46 C/C genotype for the null phenotype were almost identical (63.4% versus 63.5%) in the cases and controls.

The DARC null phenotype is, however, associated with neutrophil levels.[51,53,54] Approximately 25% to 50% of persons of African descent have persistent benign ethnic neutropenia, with low leukocyte and neutrophil counts below the normalized range for populations of European descent. Leukopenia with neutropenia was first described in 1941 when a physician found that black sharecroppers had lower mean white blood cell counts than a comparable group of white labors.[55] This common condition is labeled benign ethnic neutropenia and is not known to be associated with an increased risk for repeated or severe bacterial infections. Recently, it was discovered that white blood counts attributable to neutropenia are associated with *DARC* expression. Interestingly, African American individuals carrying either 1 or 2 copies of the European-type −46T allele have a normal white blood cell count range, and this indicates that the −46T allele is dominant over the African-type −46C null allele. On the other hand, individuals who inherit 2 *DARC* null alleles (−46C/C) have lower neutrophil counts. The alternative European-type (−46T) and African-type (−46C) alleles are highly predictive of the neutrophil count (Figure 1).[54]

In summary, these studies show that the *DARC* null genotype is a predictor of the benign ethnic neutropenia frequently observed in persons of African descent. On the other hand, there is no convincing evidence that the *DARC* null genotype is associated with an increased risk of HIV acquisition in people of African descent or origin.

Although studies show that the DARC null genotype is a predictor of the benign ethnic neutropenia frequently observed in persons of African descent there is no convincing evidence that the DARC null genotype is associated with an increased risk of HIV acquisition in people of African descent or origin.

HIV-ASSCOAITED NEPHROPATHY AND *MYH9*

Kidney disease represents a major health disparity for people of African origin or ancestry; the lifetime risk of ESRD is 2.5% for European Americans and 7.5% for African Americans.[56] In 1994, the National Institutes of Health FSGS study began enrolling African and European Americans with biopsy-proven FSGS and HIVAN with collapsing glomerulosclerosis to investigate the genetic correlates to these syndromes; both syndromes are more frequent in people of African ancestry.[14] Because African Americans are admixed and on average have 15% to 20% European ancestry, their chromosomes are a mosaic of European and African segments. Through the use of a panel of ancestry informative markers that are evenly distributed across the genome and have large allele frequency differences (30% on average) between Europeans and African Yoruba, the ancestral origin of chromosomal segments can be determined with statistical methods.[57,58] Admixture mapping provides a powerful method for mapping susceptibility genes for diseases/traits (eg, kidney disease) that show differential risks (odds ratio ≥2) between 2 discrete, historically isolated populations that have undergone recent admixture.

In this case, because persons of African descent are more likely to develop HIVAN and FSGS, we hypothesized that the risk variation associated with these syndromes would be located on an African-derived chromosomal segment. With a genome-wide admixture scan, *MYH9*, which encodes the molecular motor non-muscle myosin heavy chain IIA, was identified as a susceptibility gene for HIVAN as well as FSGS and nondiabetic ESRD.[59,60] Although many *MYH9* single-nucleotide polymorphisms (SNPs) were strongly associated with HIVAN and FSGS, alleles for 3 highly correlated SNPs in intron 23 (rs4821480, rs2032487, and rs4821481) and 1 in intron 14 (rs3752462) composed an extended (E) haplotype block that was more informative than any single SNP for predicting HIVAN or FSGS. Both the risk alleles and the risk E-1 haplotype that they define are extremely common in African Americans (>60%) and rare in European Americans (<4%). In our study of 58 individuals with HIVAN, 76% carried 2 copies of the risk E-1 haplotype, and 24% carried 1 copy of the risk E-1 haplotype (Figure 2). In contrast, in the hypernormal control group comprising African Americans with HIV-1 infection for 8 or more years with normal kidney function, only 35% of the participants carried 2 copies of the risk haplotype. Recent studies indicate that *MYH9* risk alleles contribute to major forms of kidney disease, including diabetic ESRD and ESRD attributed to hypertension.[61,62] However, we do not yet understand the mechanism leading to *MYH9* renal disease. To date, no coding region mutations have been found that account for the associations. We speculate that 1 or more noncoding SNPs alter expression or slice patterns.

The *MYH9* risk alleles and E-1 haplotype are remarkable in being very frequent and having very large effect sizes. The attributable risk of HIVAN for *MYH9* E-1 haplotype carriers is 100%; attributable risk is a measure of how much of the disease would be eliminated if the exposure (eg, the MYH9 risk factor) were not present. These findings have

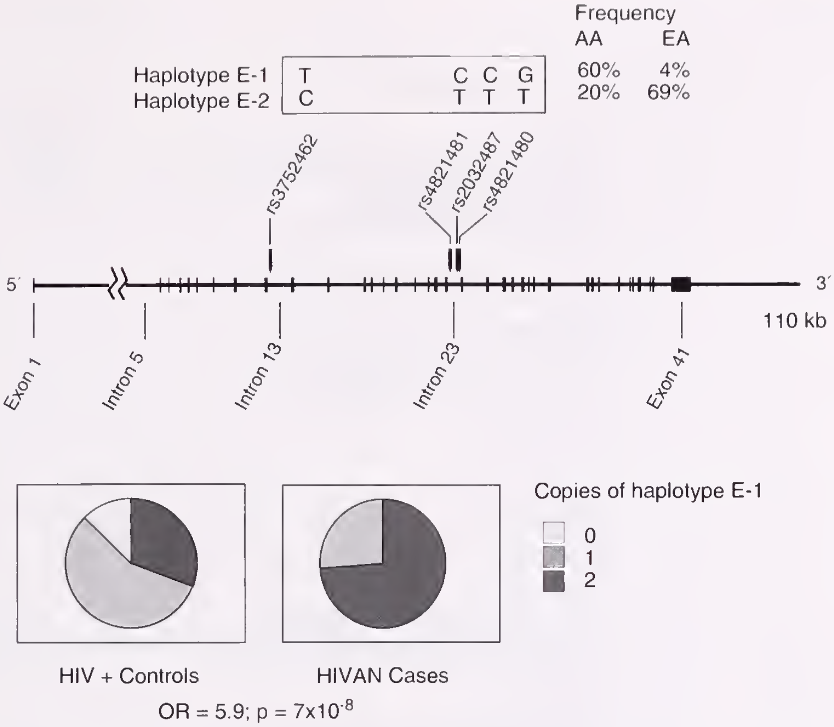


Fig 2. The E haplotype block with the alleles indicated for the most common E haplotypes in Africans (E-1) and Europeans (E-2). SNPs tagging the haplotype block are indicated; however, using only SNPs rs4821481 and rs3732462 is sufficient for inferring the E-1 haplotype. The frequencies of the 3 genotypes for the *MYH9* E-1 haplotype (rs4821480, rs2032487, rs4821481, and rs3752462), the most frequent haplotype containing the 4 SNPs, are shown for AAs with biopsy-proven HIVAN cases and hypernormal controls infected with HIV for 8 or more years with normal kidney function. Odds ratios and *P* values for the recessive model comparing focal and segmental glomerulosclerosis cases and controls are shown. **Abbreviations:** AA, African American; E, extended; EA, European American; HIV, human immunodeficiency virus; HIVAN, human immunodeficiency virus–associated nephropathy; OR, odds ratio; SNP, single-nucleotide polymorphism.

been replicated and provide strong evidence that *MYH9* genetic variation contributes to the excess burden of not only HIVAN but also other major forms of kidney disease in individuals of African descent; they also provide a genetic basis for a major US and global health disparity.[58,59,60,61]

> *These findings provide strong evidence that MYH9 genetic variation contributes to the excess burden of not only HIVAN but also other major forms of kidney disease in individuals of African descent; they also provide a genetic basis for a major US and global health disparity.*

GENES THAT MODIFY HUMAN IMMUNODEFICIENCY VIRUS

Since the discovery of CCR5 Δ32 over a decade ago, a growing number of genes have been identified that modify HIV/AIDS. These genes and their effects have been reviewed extensively elsewhere.[4,6] Many of the genes that are predicted to influence HIV acquisition are immune response genes [e.g., human leukocyte antigen (*HLA*)], intrinsic or innate antiretroviral genes [e.g., killer immunoglobin-like receptors (*KIR*)], or members of the chemokine receptor nexus. Figure 3 provides the population frequencies for the minor alleles for many of the SNPs or genetic variants that are reported to affect HIV progression, HIV levels, or HIV acquisition; data from the International HapMap Project and the Single Nucleotide Polymorphism Database of the National Center for Biotechnology Information

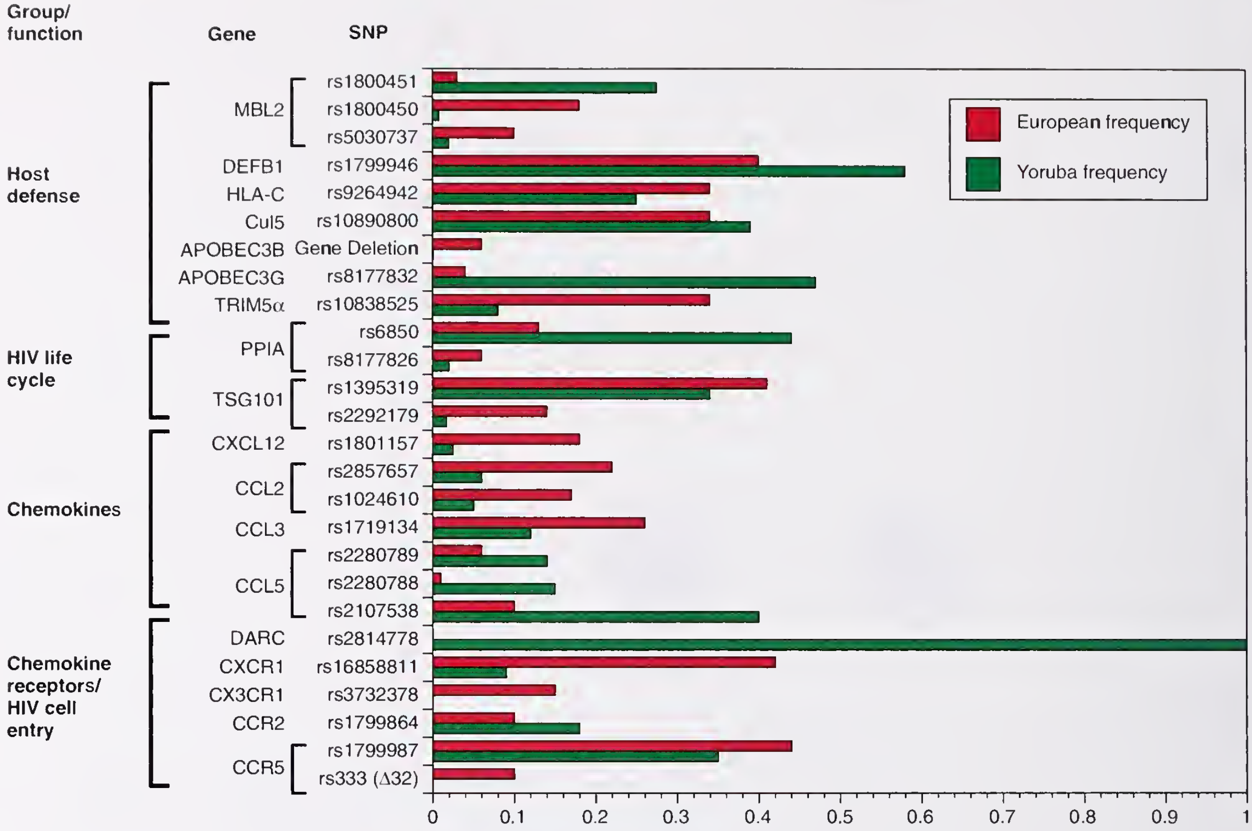


Fig 3. Allele frequencies for gene polymorphisms associated with HIV-1 disease[6] in Yoruba from Ibadan, Nigeria and in northern Europeans from Utah. HapMap frequencies were available from the Single Nucleotide Polymorphism Database of the National Center for Biotechnology Information. **Abbreviations:** APOBEC3, apolipoprotein B messenger RNA–editing enzyme catalytic polypeptide 3; CCL, chemokine (C-C motif) ligand; CCR, chemokine (C-C motif) receptor; Cul5, cullin 5; CXCL12, chemokine (C-X-C motif) ligand 12; CX3CR1, chemokine (C-X3-C motif) receptor 1; CXCR1, cysteine-X-cysteine receptor 1; DARC, Duffy antigen receptor for chemokines; DEFB1, defensin beta 1; HIV, human immunodeficiency virus; HLA-C, human leukocyte antigen C; MBL2, mannose-binding lectin (protein C) 2; PPIA, peptidylprolyl isomerase A (cyclophilin A); TRIM5α, tripartite motif-containing 5 alpha; TSG101, tumor susceptibility gene 101.

have been used to provide allele frequencies for West Africans (Yoruba from Ibadan, Nigeria) and northern Europeans from Utah. The allele frequencies for these 2 groups are representative of allele frequencies observed in African Americans and European Americans. It is immediately apparent that the majority of these polymorphisms that have been implicated in HIV pathogenesis show quite different allele frequencies between the 2 continental groups. Despite these differences in allele frequencies between African and European groups, there appears to be no major genetic predisposition to faster progression to AIDS between individuals of European descent and individuals of African descent. A comparison of self-reporting African American injecting drug users from Baltimore and European American men who have sex with men from 4 US cities indicate no significance differences in AIDS-free survival (Winkler, unpublished data) between these disparate groups with vastly different access to health care and lifestyles. Similarly, mother-to-infant transmission rates are similar among ethnic/racial groups, and this suggests that there is no major population-specific genetic predictor of HIV transmission and acquisition.[63]

Despite these differences in allele frequencies between African and European groups, there appears to be no major genetic predisposition to faster progression to AIDS between individuals of European descent and individuals of African descent.

CONCLUSION

Since the discovery of CCR5 Δ32 as a major determinant of resistance to HIV infection, considerable effort has been expended to identify genetic variation that predicts infection, pathogenesis, and disease outcomes. The discovery of such genes may provide needed insights into pathways used by the virus for completion of its lifecycle and provide targets for therapeutic interventions. Genetic dissection of host defense genes may also point the way to more effective vaccines or biological modifiers to augment natural immune surveillance. However, it is essential that well-designed studies be carried out in diverse populations because people of color carry much of the burden of HIV/AIDS throughout the world. An understanding of the genetic correlates of HIV disease in all ethnic/racial groups is crucial if the goals of personalized medicine, improved therapeutic interventions, and informed public health policies are to be realized.

ACKNOWLEDGMENT

The author thanks Dr. George Nelson, Dr. Bert Gold, Dr. Ping An, and Dr. Jeffrey Kopp for insightful discussions and the participants who made these studies possible. This project has been funded in whole or in part with federal funds from the National Cancer Institute (National Institutes of Health) under contract HHSN26120080001E. This research was supported by the Intramural Research Program of the Center for Cancer Research (National Cancer Institute, National Institutes of Health).
The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. *2008 Report on the Global AIDS Epidemic*. Geneva, Switzerland: Joint United Nations Programme on HIV/AIDS; 2008.
2. Centers for Disease Control and Prevention. Cases of HIV infection and AIDS in the United States and dependent areas, 2007. http://www.cdc.gov/hiv/topics/surveillance/resources/reports/2007report/default.htm. Accessed January 2010.
3. Sutton MY, Jones RL, Wolitski RJ, et al. A review of the Centers for Disease Control and Prevention's response to the HIV/AIDS crisis among blacks in the United States, 1981–2009. *Am J Public Health* 2009; 99(suppl 2): S351–S359.
4. O'Brien SJ, Nelson GW. Human genes that limit AIDS. *Nat Genet* 2004; 36: 565–574.
5. Fellay J, Shianna KV, Ge D, et al. A whole-genome association study of major determinants for host control of HIV-1. *Science* 2007; 317: 944–947.
6. Fellay J. Host genetics influences on HIV type-1 disease. *Antivir Ther* 2009; 14: 731–738.
7. Le Clerc S, Limou S, Coulonges C, et al. Genomewide association study of a rapid progression cohort identifies new susceptibility alleles for AIDS (ANRS Genomewide Association Study 03). *J Infect Dis* 2009; 200: 1194–1201.
8. Limou S, Le Clerc S, Coulonges C, et al. Genomewide association study of an AIDS-nonprogression cohort emphasizes the role played by HLA genes (ANRS Genomewide Association Study 02). *J Infect Dis* 2009; 199: 419–426.
9. Dalmasso C, Carpentier W, Meyer L, et al. Distinct genetic loci control plasma HIV-RNA and cellular HIV-DNA levels in HIV-1 infection: the ANRS Genome Wide Association 01 study. *PLoS One* 2008; 3: e3907.
10. Rao TK, Filippone EJ, Nicastri AD, et al. Associated focal and segmental glomerulosclerosis in the acquired immunodeficiency syndrome. *N Engl J Med* 1984; 310: 669–673.
11. Cantor ES, Kimmel PL, Bosch JP. Effect of race on expression of acquired immunodeficiency syndrome-associated nephropathy. *Arch Intern Med* 1991; 151: 125–128.
12. Praditpornsilpa K, Napathorn S, Yenrudi S, et al. Renal pathology and HIV infection in Thailand. *Am J Kidney Dis* 1999; 33: 282–286.
13. Winston JA, Burns GC, Klotman PE. The human immunodeficiency virus (HIV) epidemic and HIV-associated nephropathy. *Semin Nephrol* 1998; 18: 373–377.
14. Kopp JB, Winkler C. HIV-associated nephropathy in African Americans. *Kidney Int Suppl* 2003; 83: S43–S49.
15. Lopes GS, Marques LP, Rioja LS, et al. Glomerular disease and human immunodeficiency virus infection in Brazil. *Am J Nephrol* 1992; 12: 281–287.
16. Nochy D, Glotz D, Dosquet P, et al. Renal disease associated with HIV infection: a multicentric study of 60 patients from Paris hospitals. *Nephrol Dial Transplant* 1993; 8: 11–19.
17. Pakasa M, Mangani N, Dikassa L. Focal and segmental glomerulosclerosis in nephrotic syndrome: a new profile of adult nephrotic syndrome in Zaire. *Mod Pathol* 1993; 6: 125–128.
18. Fabian J, Naicker S. HIV and kidney disease in sub-Saharan Africa. *Nat Rev Nephrol* 2009; 5: 591–598.
19. Naicker S, Han TM, Fabian J. HIV/AIDS–dominant player in chronic kidney disease. *Ethn Dis* 2006; 16: S2-56–60.
20. Satko SG, Sedor JR, Iyengar SK, Freedman BI. Familial clustering of chronic kidney disease. *Semin Dial* 2007; 20: 229–236.
21. Coop G, Pickrell JK, Novembre J, et al. The role of geography in human adaptation. *PLoS Genet* 2009; 5: e1000500.

22. Hancock AM, Witonsky DB, Gordon AS, et al. Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet* 2008; 4: e32.
23. Campbell MC, Tishkoff SA. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annu Rev Genomics Hum Genet* 2008; 9: 403–433.
24. Reed FA, Tishkoff SA. African human diversity, origins and migrations. *Curr Opin Genet Dev* 2006; 16: 597–605.
25. Tishkoff SA, Kidd KK. Implications of biogeography of human populations for 'race' and medicine. *Nat Genet* 2004; 36: S21–S7.
26. Dean M, Carrington M, Winkler C, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* 1996; 273: 1856–1862.
27. Huang Y, Paxton WA, Wolinsky SM, et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat Med* 1996; 2: 1240–1243.
28. Liu R, Paxton WA, Choe S, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996; 86: 367–377.
29. McDermott DH, Beecroft MJ, Kleeberger CA, et al. Chemokine RANTES promoter polymorphism affects risk of both HIV infection and disease progression in the Multicenter AIDS Cohort Study. *AIDS* 2000; 14: 2671–2678.
30. O'Brien TR, Winkler C, Dean M, et al. HIV-1 infection in a man homozygous for CCR5 delta 32. *Lancet* 1997; 349: 1219.
31. Stephens JC, Reich DE, Goldstein DB, et al. Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am J Hum Genet* 1998; 62: 1507–1515.
32. Martinson JJ, Chapman NH, Rees DC, et al. Global distribution of the CCR5 gene 32-basepair deletion. *Nat Genet* 1997; 16: 100–103.
33. Galvani AP, Slatkin M. Evaluating plague and smallpox as historical selective pressures for the CCR5-Δ32 HIV-resistance allele. *Proc Natl Acad Sci U S A* 2003; 100: 15276–15279.
34. Dean M, Carrington M, Winkler C, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science* 1996; 273: 1856–1862.
35. Hendrickson SL, Jacobson LP, Nelson GW, et al. Host genetic influences on highly active antiretroviral therapy efficacy and AIDS-free survival. *J Acquir Immune Defic Syndr* 2008; 48: 263–271.
36. Ometto L, Zanchetta M, Mainardi M, et al. Co-receptor usage of HIV-1 primary isolates, viral burden, and CCR5 genotype in mother-to-child HIV-1 transmission. *AIDS* 2000; 14: 1721–1729.
37. Schliekelman P, Garner C, Slatkin M. Natural selection and resistance to HIV. *Nature* 2001; 411: 545–546.
38. Michon P, Woolley I, Wood EM, et al. Duffy-null promoter heterozygosity reduces DARC expression and abrogates adhesion of the *P. vivax* ligand required for blood-stage infection. *FEBS Lett* 2001; 495: 111–114.
39. Hill AV. Malaria resistance genes: a natural selection. *Trans R Soc Trop Med Hyg* 1992; 86: 225–226.
40. Horuk R, Chitnis CE, Darbonne WC, et al. A receptor for the malarial parasite *Plasmodium vivax*: the erythrocyte chemokine receptor. *Science* 1993; 261: 1182–1184.
41. Tournamille C, Colin Y, Cartron JP, Le Van Kim C. Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nat Genet* 1995; 10: 224–228.
42. Mayr FB, Spiel AO, Leitner JM, et al. Racial differences in endotoxin-induced tissue factor-triggered coagulation. *J Thromb Haemost* 2009; 7: 634–640.
43. Mayr FB, Spiel AO, Leitner JM, et al. Duffy antigen modifies the chemokine response in human endotoxemia. *Crit Care Med* 2008; 36: 159–165.
44. He W, Neil S, Kulkarni H, et al. Duffy antigen receptor for chemokines mediates trans-infection of HIV-1 from red blood cells to target cells and affects HIV-AIDS susceptibility. *Cell Host Microbe* 2008; 4: 52–62.
45. Lachgar A, Jaureguiberry G, Le Buenac H, et al. Binding of HIV-1 to RBCs involves the Duffy antigen receptors for chemokines (DARC). *Biomed Pharmacother* 1998; 52: 436–439.
46. McKeigue PM, Carpenter JR, Parra EJ, Shriver MD. Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: application to African-American populations. *Ann Hum Genet* 2000; 64: 171–186.
47. Hoggart CJ, Parra EJ, Shriver MD, et al. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet* 2003; 72: 1492–1504.
48. Tian C, Gregersen PK, Seldin MF. Accounting for ancestry: population substructure and genome-wide association studies. *Hum Mol Genet* 2008; 17: R143–R50.
49. Walley NM, Julg B, Dickson SP, et al. The Duffy antigen receptor for chemokines null promoter variant does not influence HIV-1 acquisition or disease progression. *Cell Host Microbe* 2009; 5: 408–410.
50. Winkler CA, An P, Johnson R, et al. Expression of Duffy antigen receptor for chemokines (DARC) has no effect on HIV-1 acquisition or progression to AIDS in African Americans. *Cell Host Microbe* 2009; 5: 411–413.
51. Julg B, Reddy S, van der Stok M, et al. Lack of Duffy antigen receptor for chemokines: no influence on HIV disease progression in an African treatment-naive population. *Cell Host Microbe* 2009; 5: 413–415.
52. Horne KC, Li X, Jacobson LP, et al. Duffy antigen polymorphisms do not alter progression of HIV in African Americans in the MACS cohort. *Cell Host Microbe* 2009; 5: 415–417.
53. Nalls MA, Wilson JG, Patterson NJ, et al. Admixture mapping of white cell count: genetic locus responsible for lower white blood cell count in the Health ABC and Jackson Heart studies. *Am J Hum Genet* 2008; 82: 81–87.
54. Reich D, Nalls MA, Kao WH, et al. Reduced neutrophil count in people of African descent is due to a regulatory variant in the Duffy antigen receptor for chemokines gene. *PLoS Genet* 2009; 5: e1000360.
55. Forbes WH, Johnson RD, Consolazio F. Leukopenia in Negro workmen. *Am J Med Sci* 1941; 201: 407–412.
56. Kiberd BA, Clase CM. Cumulative risk for developing end-stage renal disease in the US population. *J Am Soc Nephrol* 2002; 13: 1635–1644.

57. Smith MW, Patterson N, Lautenberger JA, et al. A high-density admixture map for disease gene discovery in African Americans. *Am J Hum Genet* 2004; 74: 1001–1013.
58. Patterson N, Hattangadi N, Lane B, et al. Methods for high-density admixture mapping of disease genes. *Am J Hum Genet* 2004; 74: 979–1000.
59. Kao WH, Klag MJ, Meoni LA, et al. MYH9 is associated with nondiabetic end-stage renal disease in African Americans. *Nat Genet* 2008; 40: 1185–1192.
60. Kopp JB, Smith MW, Nelson GW, et al. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet* 2008; 40: 1175–1184.
61. Freedman BI, Hicks PJ, Bostrom MA, et al. Non-muscle myosin heavy chain 9 gene MYH9 associations in African Americans with clinically diagnosed type 2 diabetes mellitus-associated ESRD. *Nephrol Dial Transplant* 2009; 24: 3366–3371.
62. Freedman BI, Hicks PJ, Bostrom MA, et al. Polymorphisms in the non-muscle myosin heavy chain 9 gene (MYH9) are strongly associated with end-stage renal disease historically attributed to hypertension in African Americans. *Kidney Int* 2009; 75: 736–745.
63. Working Group on Mother-To-Child Transmission of HIV. Rates of mother-to-child transmission of HIV-1 in Africa, America, and Europe: results from 13 perinatal studies. *J Acquir Immune Defic Syndr Hum Retrovirol* 1995; 8: 506–510.

Mount Sinai Journal of Medicine 77:160–165, 2010

Genetic and Environmental Contributions to Racial Disparities in Preterm Birth

Siobhan M. Dolan, MD, MPH

Department of Obstetrics and Gynecology and Women's Health, Albert Einstein College of Medicine, Bronx, NY

OUTLINE

Preterm Birth
Genetic Contribution to Preterm Birth
Gene-Environment Interactions
Pathophysiological Pathways to Preterm Birth
Inflammation
Maternal/Fetal Stress
Uterine Distension
Decidual Hemorrhage
Metabolic Enzymes
Conclusion

ABSTRACT

The preterm birth rate exceeds 12% in the United States, and preterm birth continues to be a clinical and public health challenge globally. Even though preterm birth is a major contributor to infant mortality and lifelong morbidity, there are few effective strategies to predict preterm birth and few clinical interventions to prevent it. Genomic research approaches that identify risk factors at the intersection of genetics and the environment will likely provide insights. Both genetic and environmental factors are known to contribute to the racial disparity seen in preterm birth. Through the identification of relevant gene-environment interactions that contribute to preterm birth and may underlie the racial disparity in preterm birth, research that will translate to clinical practice and ultimately prevent a number of preterm births is possible. ***Mt Sinai J Med 77:160–165, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: common complex disorder, genetics, genomics, preterm birth.

Address Correspondence to:

Siobhan M. Dolan
Department of Obstetrics and Gynecology and Women's Health
Albert Einstein College of Medicine
Bronx, NY
Email: siobhanmdolan@yahoo.com

The preterm birth rate exceeds 12% in the United States, and preterm birth continues to be a clinical and public health challenge globally.[1] Persistent racial disparities have been observed. In 2007, the rate of preterm birth in the United States was highest for black infants (18.3%), who were followed by Native Americans (13.9%), Hispanics (12.3%), whites (11.5%), and Asians (10.9%; Figure 1).[2]

> *In 2007, the rate of preterm birth in the United States was highest for black infants (18.3%), who were followed by Native Americans (13.9%), Hispanics (12.3%), whites (11.5%), and Asians (10.9%).*

Defined as birth before the completion of 37 weeks of gestation, preterm birth is the leading cause of infant mortality (from birth to 1 year of life) among black infants and the second leading cause among all infants.[3] In analyses that used a more inclusive classification of preterm-related causes of death, an even higher proportion of infant deaths (36.5% in 2006) were attributed to preterm birth,[4] and the preterm-related infant mortality rate for black infants was 3.5 times higher than that for whites.[5] However, despite the impact of preterm birth on birth outcomes and lifelong health, there are few effective

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20169

© 2010 Mount Sinai School of Medicine

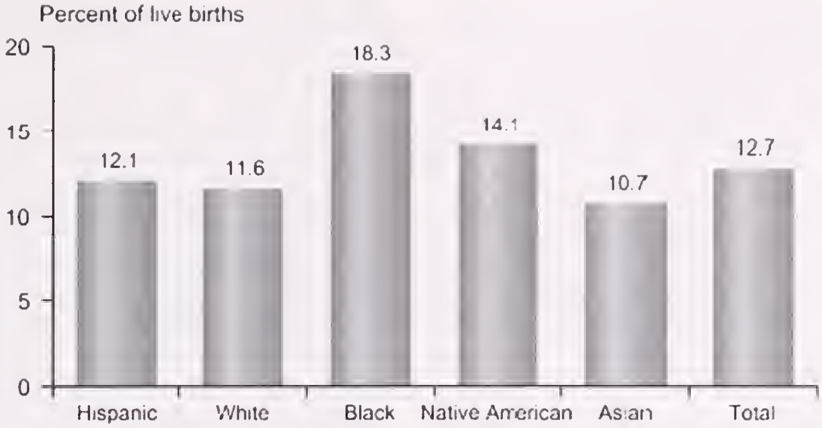


Fig 1. Preterm births by race/ethnicity: United States, 2004–2006 average. Reprinted with permission from the March of Dimes Foundation.[47]

strategies to predict or prevent it. Furthermore, the explanation for the persistent racial disparity is poorly understood. There is tremendous interest in using the new research directions that genetics and genomics can provide to better understand the etiology and racial disparities in preterm birth.

PRETERM BIRTH

Many sociodemographic risk factors for preterm birth have been identified. In the United States, racial variation in preterm birth has been consistently identified and likely reflects genetic predisposition, environmental factors including differential access to care and varying socioeconomic status, and a complex interplay of gene-environment interactions.[6,7]

> *In the United States, racial variation in preterm birth likely reflects genetic predisposition, environmental factors including differential access to care and varying socioeconomic status, and a complex interplay of gene-environment interactions.*

Other factors known to increase the risk of preterm birth include cigarette smoking,[8] extremes of maternal age (younger women and older women), multiple gestation,[9] and use of assisted reproductive technology in singletons as well as multiple births.[10] Much of the increase in the preterm birth rate over recent years can be attributed to increases in late preterm births between 34 0/7 and 36 6/7 weeks of gestation.[11]

A major challenge in studying preterm birth is consistency in the definition of preterm birth. Preterm birth following spontaneous labor may be distinct from that following preterm premature rupture of the membranes (PPROM), and thus a clearly defined phenotype is of utmost importance in performing research.[12] Among epidemiological correlates of preterm birth, one of the most consistent is that the risk of preterm birth is increased in women with a previous preterm birth.[13] Traditionally, this observation has been explained by the persistence of environmental risk factors such as poverty, stress, and limited access to prenatal care as well as social and nutritional deficits that secondarily increase the risk of infections and other medical disorders.[7,14] However, the increased recurrence risk may also be explained by genetic susceptibility that persistently regulates a woman's response to environmental stressors.

GENETIC CONTRIBUTION TO PRETERM BIRTH

In a familial aggregation study in Utah, Porter *et al.*[15] demonstrated that a component of preterm birth risk is heritable. Using linked birth cohorts composed of 1405 preterm mothers and 2781 term mothers, they showed an increased risk of preterm delivery for women who themselves were born before 37 weeks of gestation (odds ratio = 1.18, 95% confidence interval = 1.02, 1.37). Furthermore, a dose-response relationship was demonstrated between a mother's own gestational age at birth and her risk of delivering her child preterm, with mothers born earlier at increased risk of

> *Porter et al. demonstrated that a component of preterm birth risk is heritable. Using linked birth cohorts composed of 1405 preterm mothers and 2781 term mothers, they showed an increased risk of preterm delivery for women who themselves were born before 37 weeks of gestation.*

delivering their children preterm. More recent work has affirmed the maternal contribution to preterm birth and demonstrated a paternal contribution, albeit smaller, as well.[16,17]

Subsequent twin studies were conducted to determine the relative importance of genetic effects on gestational length. In a Swedish study,[18] twin concordance rates and intraclass correlation for gestational length were consistently higher in monozygotic twins than dizygotic twins, and this suggested a genetic contribution to gestational length with a heritability index for preterm birth in the range of 36% (95% confidence interval = 0.03–0.51). Similarly, a twin study from Australia revealed a heritability index ranging from 17% to 27%.[19] These studies suggest that somewhere between a quarter and a third of the variation in the risk of preterm birth is attributable to genetic influences.

In considering genetics, race is important as it not only reflects social and cultural affiliation but often can serve as a proxy for geographic ancestry. The fact that race can provide some clues about shared ancestry and therefore shared genetics is important but complicated as the relationship between race and genetics is not precise.[20–22] Furthermore, race can also be a proxy for other factors such as discrimination and inequities that have little to do with genetics but underpin racial disparities in health outcomes, including preterm birth.

> *Although race can serve as a proxy for geographic ancestry race can also be a proxy for other factors such as discrimination and inequities that have little to do with genetics but underpin racial disparities in health outcomes, including preterm birth.*

GENE-ENVIRONMENT INTERACTIONS

A gene-environment interaction is defined as "a different effect of an environmental exposure on disease risk in persons with different genotypes."[23] Evidence for the contribution of gene-environment interactions to the occurrence of low birthweight and preterm birth was put forth by Wang *et al.*[24] in 2002. Their study examined whether the association between maternal cigarette smoking and preterm birth differed with polymorphisms of 2 maternal cytochrome p450 genes, cytochrome P450 1A1 (CYP1A1) and glutathione *S*-transferase theta 1 (GSTT1). Women who smoked and had the particular polymorphisms under study delivered more than 5 weeks earlier than nonsmokers without the genetic variants. This effect was greater than the sum of the individual effects of the polymorphisms and maternal smoking. The identification of a subgroup of women who would benefit tremendously from targeted smoking cessation efforts illustrates the potential offered by the genomic approach.

PATHOPHYSIOLOGICAL PATHWAYS TO PRETERM BIRTH

The genomic approach to preterm birth suggests that a combination of common genetic variants (polymorphisms), while not explicitly causing disease, will render some women more susceptible to the environmental influences that are well known to contribute to preterm birth.

> *The genomic approach to preterm birth suggests that a combination of common genetic variants (polymorphisms), while not explicitly causing disease, will render some women more susceptible to the environmental influences that are well known to contribute to preterm birth.*

To understand where candidate polymorphisms might lie, pathophysiological pathways leading to preterm birth can be analyzed for specific genetic variants that are consistently associated with a higher risk of preterm birth. Although these pathways might overlap and intersect, they have traditionally included (1) inflammation, (2) maternal/fetal stress,

(3) uterine distension, (4) decidual hemorrhage, and (5) metabolic enzymes.[25,26]

Inflammation

Inflammatory responses are often initiated by exposure to foreign pathogens. Several studies have looked at the association between single nucleotide polymorphisms (SNPs) in the gene for the inflammatory mediator tumor necrosis factor α (TNF-α)[27,28] and preterm birth and the role that race might play. TNF-α is part of the major histocompatibility complex on chromosome 6p, and it has a promoter region at -308 bp that has been shown to be highly polymorphic in the population. Dizon-Townson *et al.*[27] used a case control study design to show that a variant in the promoter region, TNF T2, which increases transcription of the gene, was not more frequent in women who delivered preterm versus women who delivered at term. In a case control study restricted to African American women who experienced a previous spontaneous preterm birth (one that followed preterm labor or PPROM), Roberts *et al.*[28] found that carriage of 1 TNF T2 polymorphism was a risk factor for PPROM (odds ratio = 3.18, 95% confidence interval = 1.33–7.83, $P = 0.008$). Another study provided preliminary data suggesting that TNF-α modifies susceptibility to bacterial vaginosis and that race/ethnicity modifies the relationship between TNF-α and preterm birth.[29] Finally, a meta-analysis of the association between the SNP in the TNF-α gene (G-308>A) and preterm birth failed to show a statistically significant association.[30] This illustrates the challenge of identifying robust gene-disease associations.[31] However, this line of research continues to be productive as Menon *et al.*[32] demonstrated higher TNF-α concentrations in the amniotic fluid of black women delivering preterm in comparison with whites, and they have performed genetic studies suggesting that race may play a key role in a differential inflammatory response that contributes to preterm birth, especially among blacks.[33–35]

Maternal/Fetal Stress

The maternal/fetal stress pathway to preterm birth has identified a key role for stress hormones, including corticotropin-releasing hormone.[36] Corticotropin-releasing hormone, which is highly polymorphic in various populations and whose levels have been correlated with pregnancy outcome,[37,38] is a candidate for further study of polymorphisms. Recent research has found that high maternal urinary catecholamine levels in women who go on to have a preterm birth may suggest excess stress and/or a predisposition to elevated sympathetic activation.[39] The complex interplay between race and stress that contributes to preterm birth continues to be the subject of promising research.

Uterine Distension

Uterine factors in abnormal uterine distension and cervical function may also be influenced by genetic variation. Matrix metalloproteinases, which affect tissue remodeling, are candidates for polymorphic variation in women who deliver preterm versus those who deliver at term.[40]

Matrix metalloproteinases play a role in PPROM, which often precedes preterm birth.[41] African American women have been found to be at increased risk for PPROM, and thus PPROM may factor into the higher rate of preterm birth for black women. In an interesting line of research, Wang *et al.*[42] identified a functional SNP in the promoter of the serpin peptidase inhibitor clade H member 1 (SERPINH1) gene that had a greater frequency in African Americans versus European Americans. The variant in SERPINH1, which plays a role in collagen synthesis, was associated with PPROM in a case control study of preterm birth, and this supports the hypothesis that genetic variation explains some of the racial disparity seen in preterm birth.

Decidual Hemorrhage

The contribution of genetic variation to inherited thrombophilias that predispose to poor perinatal outcomes, including preterm birth and multiple miscarriages,[43] is well established. The description of variants in factor V Leiden, prothrombin G20210A, methylenetetrahydrofolate reductase, and the type 1 plasminogen activator inhibitor gene[44] has already established the scientific basis for some clinical interventions in the form of anticoagulation[45] and has allowed the initiation of studies designed to evaluate women with preterm birth with respect to these variants. Such studies will also need to consider the interaction between these genetic variants and race/ethnicity as they contribute to preterm birth.

Metabolic Enzymes

As cited previously, Wang *et al.*[24] reported 2 polymorphic maternal metabolic genes, CYP1A1 and GSTT1, that are part of the cytochrome p450 system and contribute to the detoxification of smoking byproducts. Certainly, there are many metabolic enzymes that are essential to achieving homeostasis. There is still the challenge of identifying additional relevant enzymes that are polymorphic in the population and that increase susceptibility to preterm birth by participating in key gene-environment interactions, which may contribute to the racial disparity seen in preterm birth.

CONCLUSION

The prevention of preterm birth will require a reevaluation of clinical practice, a focus on disparities, and a comprehensive research strategy using genomic approaches. Screening tools, including questions about the personal and family history of preterm birth, will allow targeted risk factor reduction strategies for high-risk women, including African American women who are at increased risk. In practical terms, a woman with a positive personal or family history of preterm birth should be offered preconception care that includes targeted smoking cessation intervention, screening for infection, and consideration of progesterone supplementation. An increased understanding of the racial disparity seen in preterm birth and the identification of genetic contributions to preterm birth as well as environmental influences and gene-environment interactions will allow the tools and promise of the genomic era to open research and therapeutic opportunities to improve pregnancy outcomes.[46]

ACKNOWLEDGMENT

This article was presented on November 17, 2008, at the Mt. Sinai School of Medicine in New York City as part of the conference *Achieving Health Equity: Genes, Race, Ethnicity and the Environment.*

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Howson CP, Merialdi M, Lawn JE. *White Paper on Preterm Birth: The Global and Regional Toll.* White Plains, NY: March of Dimes Foundation; 2009.
2. Hamilton BE, Martin JA, Ventura SJ. Births: preliminary data for 2007. *Natl Vital Stat Rep* 2009; 57: 1–23.
3. Mathews TJ, MacDorman MF. Infant mortality statistics from the 2005 period: linked birth/infant death data set. *Natl Vital Stat Rep* 2008; 57: 1–32.
4. Callaghan WM, MacDorman MF, Rasmussen SA, et al. The contribution of preterm birth to infant mortality rates in the United States. *Pediatrics* 2006; 118: 1566–1573.
5. MacDorman MF, Callaghan WM, Mathews TJ, et al. Trends in preterm-related infant mortality by race and ethnicity, United States, 1999–2004. *Int J Health Serv* 2007; 37: 635–641.
6. Zhang J, Savitz DA. Preterm birth subtypes among blacks and whites. *Epidemiology* 1992; 3: 428–433.
7. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. *Lancet* 2008; 371: 75–84.
8. Shah NR, Bracken MB. A systematic review and meta-analysis of prospective studies on the association between maternal cigarette smoking and preterm delivery. *Am J Obstet Gynecol* 2000; 182: 465–472.
9. Russell RB, Petrini JR, Damus K, et al. The changing epidemiology of multiple births in the United States. *Obstet Gynecol* 2003; 101: 129–135.
10. Reynolds MA, Schieve LA, Martin JA, et al. Trends in multiple births conceived using assisted reproductive technology, United States, 1997–2000. *Pediatrics* 2003; 111(pt 2): 1159–1162.
11. Engle WA, Tomashek KM, Wallman C. "Late-preterm" infants: a population at risk. *Pediatrics* 2007; 120: 1390–1401.
12. Dolan SM. Preterm birth. In: Khoury M, Bedrosian SR, Gwinn M, Higgins JPT, eds. *Human Genome Epidemiology: Building the Evidence for Using Genetic Information to Improve Health and Prevent Disease.* 2nd ed. New York, NY: Oxford University Press; 2009; 357.
13. Adams MM, Elam-Evans LD, Wilson HG, Gilbertz DA. Rates of and factors associated with recurrence of preterm delivery. *JAMA* 2000; 283: 1591–1596.
14. Iams JD. The epidemiology of preterm birth. *Clin Perinatol* 2003; 30: 651–664.
15. Porter TF, Fraser AM, Hunter CY, et al. The risk of preterm birth across generations. *Obstet Gynecol* 1997; 90: 63–67.
16. Lie RT, Wilcox AJ, Skjaerven R. Maternal and paternal influences on length of pregnancy. *Obstet Gynecol* 2006; 107: 880–885.
17. Wilcox AJ, Skaerven R, Lie RT. Familial patterns of preterm delivery: maternal and fetal contributions. *Am J Epidemiol* 2008; 167: 474–479.
18. Clausson B, Lichtenstein P, Cnattingius S. Genetic influence on birthweight and gestational length determined by studies in offspring of twins. *BJOG* 2000; 107: 375–381.
19. Treloar SA, Macones GA, Mitchell LE, Martin NG. Genetic influences on premature parturition in an Australian twin sample. *Twin Res* 2000; 3: 80–82.
20. Madhere S, Harrell J, Royal CD. Social ecology, genomics, and African American health: a nonlinear dynamical perspective. *J Black Psychol* 2009; 35: 154–179.
21. Collins FS. What we do and don't know about 'race', 'ethnicity', genetics and health at the dawn of the genome era. *Nat Genet* 2004; 36(suppl): S13–S15.

22. Anum EA, Springel EH, Shriver MD, Strauss JF III. Genetic contributions to disparities in preterm birth. *Pediatr Res* 2009; 65: 1–9.
23. Ottman R. Gene-environment interaction: definitions and study designs. *Prev Med* 1996; 25: 764–770.
24. Wang X, Zuckerman B, Pearson C, et al. Maternal cigarette smoking, metabolic gene polymorphism, and infant birth weight. *JAMA* 2002; 287: 195–202.
25. Lockwood CJ, Kuczynski E. Risk stratification and pathological mechanisms in preterm delivery. *Paediatr Perinat Epidemiol* 2001; 15(suppl 2): 78–89.
26. Hao K, Wang X, Niu T, et al. A candidate gene association study on preterm delivery: application of high-throughput genotyping technology and advanced statistical methods. *Hum Mol Genet* 2004; 13: 683–691.
27. Dizon-Townson DS, Major H, Varner M, Ward K. A promoter mutation that increases transcription of the tumor necrosis factor-alpha gene is not associated with preterm delivery. *Am J Obstet Gynecol* 1997; 177: 810–813.
28. Roberts AK, Monzon-Bordonaba F, Van Deerlin PG, et al. Association of polymorphism within the promoter of the tumor necrosis factor alpha gene with increased risk of preterm premature rupture of the fetal membranes. *Am J Obstet Gynecol* 1999; 180: 1297–1302.
29. Macones GA, Parry S, Elkousy M, et al. A polymorphism in the promoter region of TNF and bacterial vaginosis: preliminary evidence of gene-environment interaction in the etiology of spontaneous preterm birth. *Am J Obstet Gynecol* 2004; 190: 1504–1508.
30. Menon R, Merialdi M, Betràn AP, et al. Analysis of association between maternal tumor necrosis factor-(alpha) promoter polymorphism (-308), tumor necrosis factor concentration, and preterm birth. *Am J Obstet Gynecol* 2006; 195: 1240–1248.
31. Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet* 2006; 38: 3–5.
32. Menon R, Thorsen P, Vogel I, et al. Racial disparity in amniotic fluid concentrations of tumor necrosis factor (TNF)-alpha and soluble TNF receptors in spontaneous preterm birth. *Am J Obstet Gynecol* 2008; 198: 533.
33. Menon R, Velez DR, Morgan N, et al. Genetic regulation of amniotic fluid TNF-alpha and soluble TNF receptor concentrations affected by race and preterm birth. *Hum Genet* 2008; 124: 243–253.
34. Velez DR, Fortunato S, Thorsen P, et al. Spontaneous preterm birth in African Americans is associated with infection and inflammatory response gene variants. *Am J Obstet Gynecol* 2009; 200: 209.
35. Menon R, Pearce B, Velez DR, et al. Racial disparity in pathophysiologic pathways of preterm birth based on genetic variants. *Reprod Biol Endocrinol* 2009; 7: 62.
36. Hobel CJ, Dunkel-Schetter C, Roesch SC, et al. Maternal plasma corticotropin-releasing hormone associated with stress at 20 weeks' gestation in pregnancies ending in preterm delivery. *Am J Obstet Gynecol* 1999; 180(pt 3): S257–S263.
37. Rich-Edwards J, Krieger N, Majzoub J, et al. Maternal experiences of racism and violence as predictors of preterm birth: rationale and study design. *Paediatr Perinat Epidemiol* 2001; 15(suppl 2): 124–135.
38. Mancuso RA, Schetter CD, Rini CM, et al. Maternal prenatal anxiety and corticotropin-releasing hormone associated with timing of delivery. *Psychosom Med* 2004; 66: 762–769.
39. Holzman C, Senagore P, Tian Y, et al. Maternal catecholamine levels in midpregnancy and risk of preterm delivery. *Am J Epidemiol* 2009; 170: 1014–1024.
40. Fortunato SJ, Menon R. Screening of novel matrix metalloproteinases (MMPs) in human fetal membranes. *J Assist Reprod Genet* 2002; 19: 483–486.
41. Wang H, Ogawa M, Wood JR, et al. Genetic and epigenetic mechanisms combine to control MMP1 expression and its association with preterm premature rupture of membranes. *Hum Mol Genet* 2008; 17: 1087–1096.
42. Wang H, Parry S, Macones G, et al. A functional SNP in the promoter of the SERPINH1 gene increases risk of preterm premature rupture of membranes in African Americans. *Proc Natl Acad Sci U S A* 2006; 103: 13463–13467.
43. Kovalevsky G, Gracia CR, Berlin JA, et al. Evaluation of the association between hereditary thrombophilias and recurrent pregnancy loss: a meta-analysis. *Arch Intern Med* 2004; 164: 558–563.
44. Glueck CJ, Phillips H, Cameron D, et al. The 4G/4G polymorphism of the hypofibrinolytic plasminogen activator inhibitor type 1 gene: an independent risk factor for serious pregnancy complications. *Metabolism* 2000; 49: 845–852.
45. Lockwood CJ. Inherited thrombophilias in pregnant patients: detection and treatment paradigm. *Obstet Gynecol* 2002; 99: 333–341.
46. Burke W. Genomics as a probe for disease biology. *N Engl J Med* 2003; 349: 969–974.
47. National Center for Health Statistics. Final natality data. http://www.marchofdimes.com/peristats. Accessed January 2010.

Mount Sinai Journal of Medicine 77:166–177, 2010

Access to Health and Health Care: How Race and Ethnicity Matter

Lynne D. Richardson, MD, FACEP, and Marlaina Norris, MD

Mount Sinai School of Medicine, New York, NY

OUTLINE

Access to Health
- **Physical Environment**
- **Social Environment**
- **Racism and Discrimination**

Access to Health Care
- **Health Insurance**
- **Racial and Ethnic Disparities in Facilitators and Barriers to Health Care**

Conclusion

ABSTRACT

Racial and ethnic disparities in health are multifactorial; they reflect differences in biological vulnerability to disease as well as differences in social resources, environmental factors, and health care interventions. Understanding and intervening in health inequity require an understanding of the disparate access to all of the personal resources and environmental conditions that are needed to generate and sustain health, a set of circumstances that constitute access to health. These include access to health information, participation in health promotion and disease prevention activities, safe housing, nutritious foods, convenient exercise spaces, freedom from ambient violence, adequate social support, communities with social capital, and access to quality health care. Access to health care is facilitated by health insurance, a regular source of care, and a usual primary care provider. Various mechanisms through which access to health and access to health care are mediated by race and ethnicity are discussed; these include the built environment, social environment, residential segregation, stress, racism, and discrimination. Empirical evidence supporting the association between these factors and health inequities is also reviewed. ***Mt Sinai J Med 77:166–177, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: access to health care, disparities in health care, ethnic disparities, health disparities, institutional racism, racial disparities, residential segregation, social determinants of health.

There is a great deal of data documenting pervasive and persistent disparities in health for certain racial and ethnic groups in this country[1–5] (see Figure 1). It is important to point out that racial and ethnic disparities in health cannot be explained away by between-group differences in socioeconomic status. In the United States, there is a significant relationship between race/ethnicity and socioeconomic status. However, there are residual racial/ethnic differences in health at every level of socioeconomic status for multiple indicators of health status, including self-rated health, heart disease mortality, hypertension, and obesity.[6,7] Race, ethnicity and socioeconomic status exert independent and interactive influences on health.

The causes of racial and ethnic disparities in health are multifactorial; they reflect differences in biological vulnerability to disease as well as differences in social resources, environmental conditions, and health care interventions.[8,9] Understanding

The causes of racial and ethnic disparities in health are multifactorial; they reflect differences in biological vulnerability to disease as well as differences in social resources, environmental conditions, and health care interventions.

Address Correspondence to:

Lynne D. Richardson, MD
Mount Sinai School of Medicine
New York, NY
Email: lynne.richardson@mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20174

© 2010 Mount Sinai School of Medicine

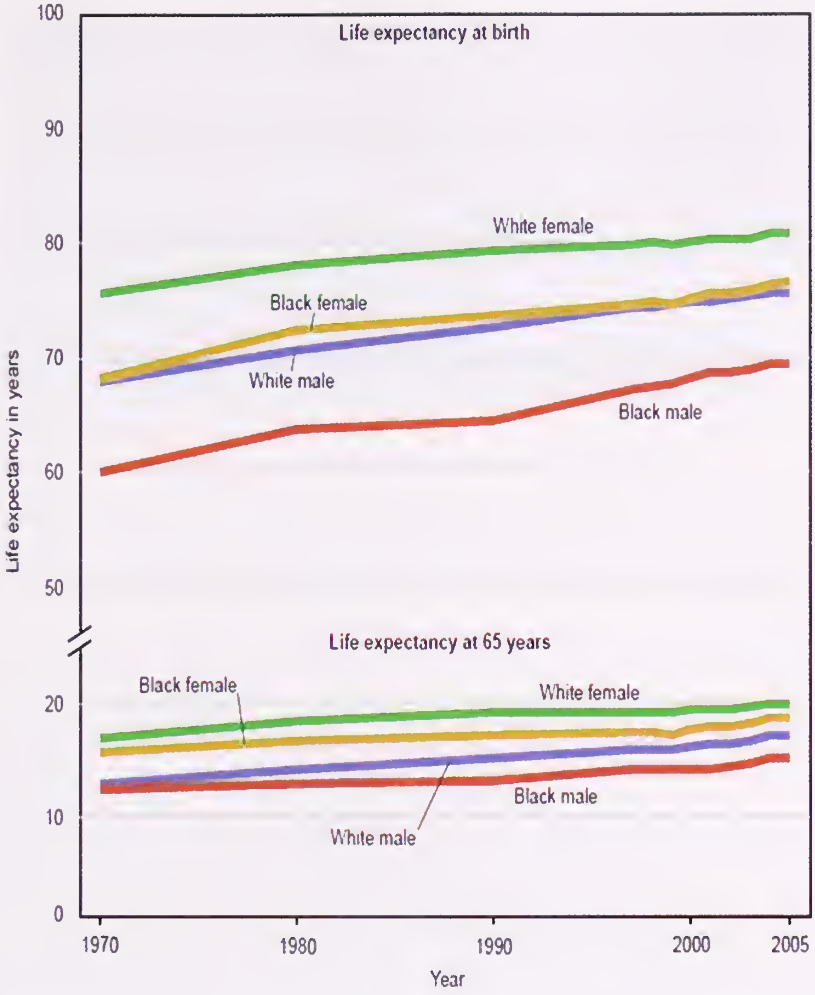


Fig 1. Life expectancy at birth and at the age of 65 years by race and sex: United States, 1970–2005. Adapted from *Health, United States, 2008 with Chartbook*.[1]

disparities requires the examination of multiple, complex, and sometimes subtle relationships between genetic susceptibility, individual behavior, social environment, physical surroundings, disease prevention, and medical treatment that lead to the observed differences in health status and health outcomes.[9]

Although major differences in the quality of care received by certain racial and ethnic minority groups have been well documented,[1,2,10,11] disparate access to health care has been shown to be a major contributor to health disparities[11–13] (see Figure 2). Access to health care is often rather narrowly construed as having health insurance, although there are many nonfinancial barriers to access at the individual, provider, institutional, and systemic levels.[14,15] Although universal health coverage is an important and even essential goal, it will not provide universal access to quality health care. Many insured individuals do not have a primary care provider or even a regular source of care and do not receive necessary preventive and therapeutic interventions. Such inadequate access to necessary health care is often, in part, a function of race and ethnicity.[10,11,16]

However, it is well known that many factors other than health care have a significant impact on health status and outcomes, and these include social determinants of health such as education and income.[8,9] Understanding and intervening in health inequity require an understanding of the disparate

Understanding and intervening in health inequity require an understanding of the disparate access to all of the resources and conditions that are needed to generate and sustain health.

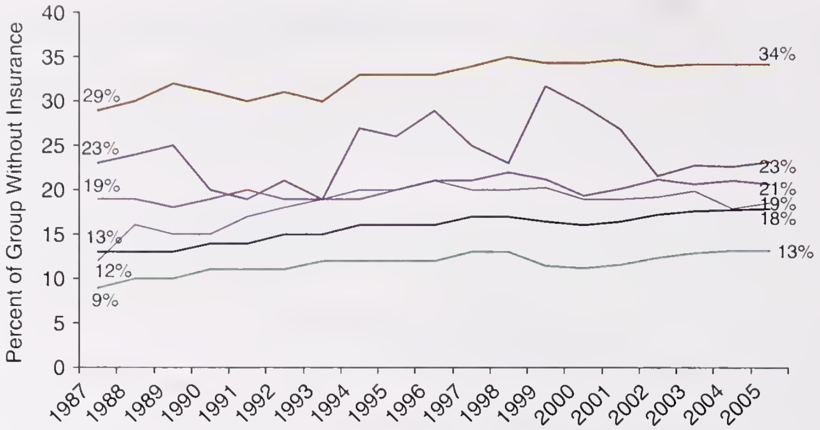


Fig 2. Uninsured under the age of 65 years by race/ethnicity: 1987–2005. The red line indicates Hispanics, the dark blue line indicates American Indians, the pink line indicates blacks, the light blue line indicates Asians, the green line indicates whites, and the black line indicates the total. Whites include all other unnamed groups. Adapted from *An Overview of the U.S. Health Care System: Chart Book*.[57]

Table 1. *Resources That Generate and Sustain Health.*

Nutritious foods	Quality healthcare
Exercise spaces	Social support
Health information	Safe housing
Health promotion	Risk-free neighborhoods
Disease prevention	Social capital

access to all of the resources and conditions that are needed to generate and sustain health, a set of circumstances that constitute access to health (see Table 1).

As the science of genomics elucidates the mechanisms by which genetic factors influence health and by which genetic susceptibility interacts with behavioral, nutritional, pharmacological, and environmental variables, it is also essential to understand how differential exposures to various behavioral, nutritional, pharmacological, and environmental variables contribute to health disparities. This article examines how race and ethnicity affect access to health and access to health care.

ACCESS TO HEALTH

Physical Environment

Where you live has a huge impact on your health because of both the built environment and the social environment. Despite the dismantling of legal segregation in the United States during the last century, demographic analyses of American communities continue to demonstrate high levels of residential segregation along racial lines.[17–19] The dissimilarity index measures the percentage of a population that would have to move to have groups evenly distributed throughout a specified geographic entity.[20] For example, under apartheid, South Africa's black-white dissimilarity index was almost 97%. Many of the largest urban areas in the United States approach that level of residential segregation (see Figure 3). The relevance of this sociodemographic phenomenon to the issue of access to health is clear. There are data relating the level of segregation to rates of morbidity and mortality; specifically, residential segregation is related to elevated risks of cause-specific and overall adult mortality,[21–24] infant mortality,[25–27] and tuberculosis.[28]

Racial and ethnic minorities disproportionately live in physical environments that lack the resources necessary to generate and sustain health. It has been shown, for example, that segregated African American neighborhoods are characterized by substandard housing, high levels of abandoned buildings and grounds, larger numbers of commercial and industrial facilities, and inadequate municipal services and amenities, including police and fire protection.[29,30] Physical locations vary with respect to exposures

The built environment may present differential risks from chemical substances such as lead paint and respiratory irritants such as diesel fumes as well as crowding, litter, and noise.

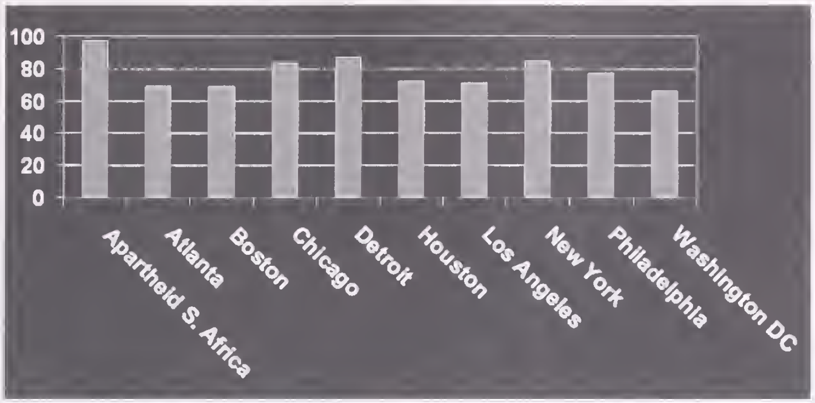


Fig 3. Residential segregation in the United States: the white-black dissimilarity index. Adapted from the CensusScope Web site.[20]

to toxins, pathogens, and carcinogens.[31] The built environment may present differential risks from chemical substances such as lead paint and respiratory irritants such as diesel fumes as well as crowding, litter, and noise.

Crowding, substandard housing, elevated noise levels, and elevated exposure to noxious pollutants and allergens (including lead, smog, particulates, and dust mites) are all common in poor, segregated communities and have all been shown to adversely affect health.[21,22] In many urban areas, the built environment may constrain exercise. Neighborhoods that lack safe recreational spaces and in which residents have concerns about their personal safety discourage individuals from participating in physical activity.[32]

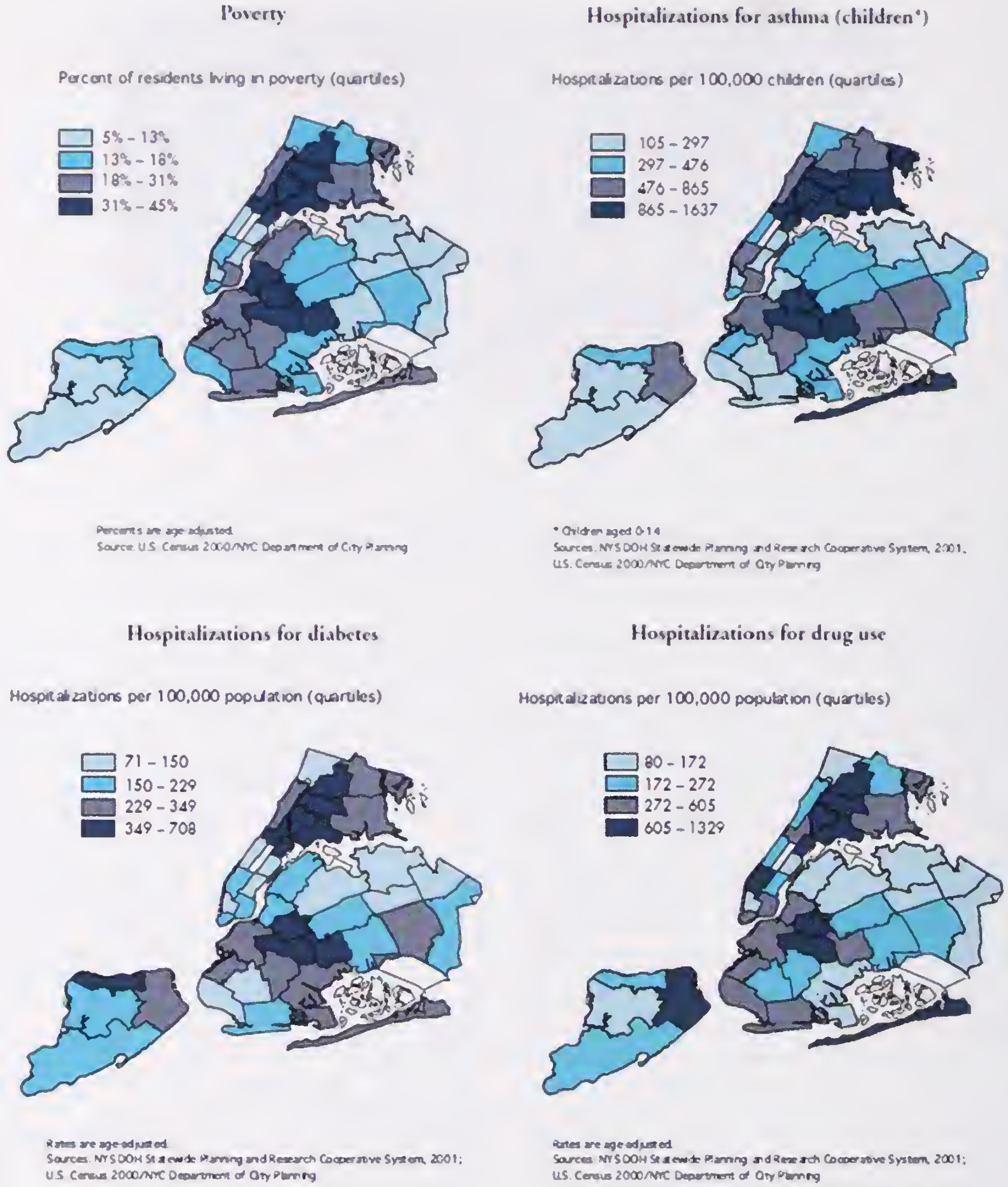
Where you live can also limit your access to healthy foods. Many large commercial enterprises avoid segregated urban areas; as a result, the available services are typically fewer in quantity, poorer in quality, and often higher in price than those available in less segregated urban and suburban areas. The absence of stores with reasonably priced fresh fruits and vegetables may discourage individuals from including these in their daily diet. The consumption of healthful foods has been shown to be positively associated with their availability, and the availability of healthful products in grocery stores varies across zip codes.[33] The high cost and poor quality of food choices in segregated neighborhoods can lead to poorer nutrition.

The absence of stores with reasonably priced fresh fruits and vegetables may discourage individuals from including these in their daily diet.

Communities with high rates of violence, tobacco use, alcohol consumption, and drug abuse place individuals at increased risk of adverse outcomes and may even facilitate risky behaviors. Many researchers have noted that both the tobacco and alcohol industries have heavily targeted poor minority communities with high-intensity billboard advertising and have concentrated alcohol and tobacco ads in print outlets with large minority readerships; researchers have also noted the increasing level of corporate sponsorship of athletic, cultural, civic, and entertainment events targeted to minority consumers.[34–38] Residential segregation may also explain racial differences in health outcomes that have strong environmental components such as homicide and drug abuse. African Americans are much more likely than whites to be victims of all types of crime, including homicide: of the 15 leading causes of death in the United States, the black-white gap is largest for homicide.[17,39] A study using national data revealed that elevated rates of cocaine use by blacks and Hispanics could be completely explained by neighborhood clusters.[40] Thus, living in certain neighborhoods places individuals at increased risks from tobacco, alcohol, drugs, and violence (see Figure 4).

Social Environment

Racial and ethnic minorities also disproportionately live in social environments that lack the resources necessary to generate and sustain health. Both health-promoting and risky behaviors may be rooted in cultural norms and are influenced by socialization within families early in the course of life. Behaviors are also influenced by education, both formal



Poverty

Percent of residents living in poverty (quartiles)

Percents are age-adjusted.
Source: U.S. Census 2000/NYC Department of City Planning

Hospitalizations for asthma (children*)

Hospitalizations per 100,000 children (quartiles)

* Children aged 0-14
Sources: NYS DOH Statewide Planning and Research Cooperative System, 2001;
U.S. Census 2000/NYC Department of City Planning

Hospitalizations for diabetes

Hospitalizations per 100,000 population (quartiles)

Rates are age-adjusted.
Sources: NYS DOH Statewide Planning and Research Cooperative System, 2001;
U.S. Census 2000/NYC Department of City Planning

Hospitalizations for drug use

Hospitalizations per 100,000 population (quartiles)

Rates are age-adjusted.
Sources: NYS DOH Statewide Planning and Research Cooperative System, 2001;
U.S. Census 2000/NYC Department of City Planning

Fig 4. Geography of race, ethnicity, poverty, and health in New York City. In New York City, African Americans and Hispanics disproportionately live in poorer neighborhoods; these neighborhoods have higher rates of morbidity and mortality for many conditions, including asthma, diabetes, human immunodeficiency virus/acquired immune deficiency syndrome, drug use, infant death, and low birthweight. Reprinted with permission from the New York City Department of Health and Mental Hygiene.[67]

> *Both health-promoting and risky behaviors may be rooted in cultural norms and are influenced by socialization within families early in the course of life.*

and experiential, and, as discussed previously, may be facilitated or discouraged by environmental factors. At the individual level, the social environment is described by factors such as social connectedness and social support.[41] These factors may provide salutary effects on health through a variety of mechanisms: mobilizing resources when they are needed to deal with problems and threats, buffering the effects of stress exposure, and accessing greater knowledge from others about health issues.[42]

At the aggregate level, lower morbidity and mortality are evident in communities with greater social capital, a construct that may be defined as the resources available to individuals and groups within communities as a result of their social network of connections.[43] A related concept is collective efficacy, which refers to the ability of community residents to undertake collective action for mutual benefit.[43] Collective efficacy and social capital have been shown to contribute to health independently of the socioeconomic status of a neighborhood.[42,43]

> *Collective efficacy and social capital have been shown to contribute to health independently of the socioeconomic status of a neighborhood.*

A growing number of studies using multilevel analyses have demonstrated an association between the social and economic characteristics of residential areas and a broad range of health outcomes that is independent of individual indicators of socioeconomic status.[44] For example, Diez Roux and colleagues[45] found that even after adjustments for education, income, occupational status, and a broad range of biomedical and behavioral risk factors for coronary heart disease (including smoking, exercise, hypertension, diabetes, obesity, and cholesterol), people residing in disadvantaged neighborhoods had a higher incidence of heart disease than people who lived in more advantaged neighborhoods. Stress and resultant behavioral and biological responses put individuals at risk for a range of diseases. Differential stress exposure may be an important mechanism by which social disadvantage gets "into the body" to affect health.[42]

Racism and Discrimination

Social scientists define racism as an organized system that categorizes population groups into races and uses this ranking to preferentially allocate societal goods and resources to groups regarded as superior.[46] This system of structuring opportunity and assigning value on the basis of the social interpretation of how we look (race) unfairly disadvantages some individuals and communities and unfairly advantages other individuals and communities. Racism generates differential treatment of members of racial outgroups (ie, groups not valued) by both individuals and social institutions; this is called discrimination.[47] Once racism is deeply embedded in a society, discrimination can persist in institutional structures and policies even in the face of declining individual levels of racial prejudice and discrimination. Moreover, negative racial stereotypes that are deeply rooted in mainstream culture can serve as an additional source of discriminatory behavior even among persons who may not be prejudiced.[7,47]

> *Once racism is deeply embedded in a society, discrimination can persist in institutional structures and policies even in the face of declining individual levels of racial prejudice and discrimination.*

Jones[48] has developed a theoretical framework for understanding racism on 3 levels: institutionalized, personally mediated, and internalized. This framework is useful for generating and testing hypotheses about the basis of race-associated differences in health outcomes and for designing effective interventions to eliminate those differences. Institutional racism, which refers to differential access to the goods, services, and opportunities of society by race, is normative, is sometimes legalized, and manifests itself both in material conditions and in access to power. Because it is structural, there does need not to be an identifiable perpetrator; examples include differential access to quality education, sound housing, gainful employment, appropriate medical facilities, a clean environment, material resources, wealth, representation in government, and control of the media. Institutionalized racism may explain the persistent association between race and socioeconomic status in this country.[48]

In addition to the indirect effects on health of discrimination through social and economic pathways, associated community exposures, and increased barriers to quality health care, some research suggests that the experience of discrimination may itself be detrimental to health. The associations are complex and appear to demonstrate that not just exposure but also responses to the experiences have health consequences.[49] Such issues are best examined with the concept of personally mediated racism, which is defined as differential assumptions about the abilities, motives, and intent of others based on their race.[48] This includes both overt prejudice and discrimination as well as subconscious bias. The Institute of Medicine report *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care* describes how this plays out in the physician-patient relationship, but this phenomenon is ubiquitous and creates many special stressors in addition to physician disrespect, such as police brutality, shopkeeper vigilance, waiter indifference, and teacher devaluation.[10,48] A recent analysis of data from the 2004 Behavioral Risk Factor Surveillance System demonstrated that being classified by others as white is associated with large and statistically significant advantages in health status, no matter how one self-identifies.[49]

Internalized racism, which is the acceptance of the negative messages about one's own abilities and intrinsic worth by members of the stigmatized group, leads to a lack of self-esteem, which may lead to risky behaviors and failure to practice healthy behaviors.[48] Jones' theoretical model is supported by a growing body of empirical evidence from population-based studies documenting the association between perceptions of racial/ethnic discrimination and health.[50] A recent review by Williams and Mohammed[7] identified several large cross-sectional studies that found a positive association between discrimination and chronic health conditions or other self-reported indicators of ill health.

ACCESS TO HEALTH CARE

Access means having "the timely use of personal health services to achieve the best health outcomes."[15] Attaining good access to care requires 3 discrete steps: gaining entry into the health care system, getting access to sites of care at which patients can receive needed services, and finding providers who meet the needs of individual patients and with whom patients can develop a relationship based on mutual communication and trust.[11,51] A myriad of patient factors, provider factors, institutional factors, and system factors affect access to health care.[14,52]

> *Attaining good access to care requires 3 discrete steps: gaining entry into the health care system, getting access to sites of care at which patients can receive needed services, and finding providers who meet the needs of individual patients and with whom patients can develop a relationship based on mutual communication and trust.*

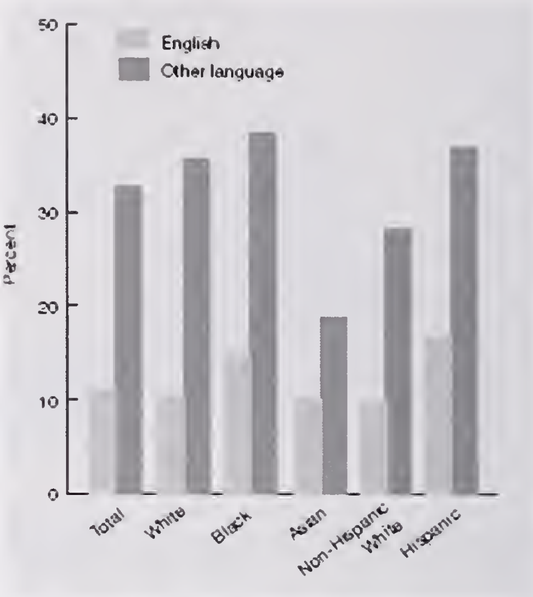


Fig 5. Uninsured by race, ethnicity, and language spoken: 2005. Reprinted with permission from the Agency for Healthcare Quality and Research.[58]

Health Insurance

In the United States, lack of health insurance is a major barrier to accessing preventive, diagnostic, or therapeutic health services. Because of the relationship between employment, income, and insurance, members of racial and ethnic minority groups who have suffered discrimination with respect to educational, economic, and credentialed immigration opportunities are more likely to be uninsured, particularly if they have limited proficiency in English (see Figure 5). Compared to those who are insured, individuals who lack health insurance are 3.9 times more likely (35% versus 9%) to not obtain needed care, 3.25 times more likely (13% versus 4%) to not fill a prescription because of cost, 4.7 times more likely (42% versus 9%) to have no regular source of care, and 3.1 times more likely (47% versus 15%) to postpone care because of cost.[53] The uninsured are more likely to die early and to have poor health status[54]; they are diagnosed at later stages of disease and get less treatment than those with insurance.[55] They are sicker when hospitalized and are more likely to die during their stay.[56]

Insurance coverage, income, and other factors explain a portion of racial and ethnic disparities in access to health care, but gaps remain after we account for these factors.[16] Since 2003, the Agency for Healthcare Research and Quality has published an annual report on health care disparities that tracks racial/ethnic disparities in access and

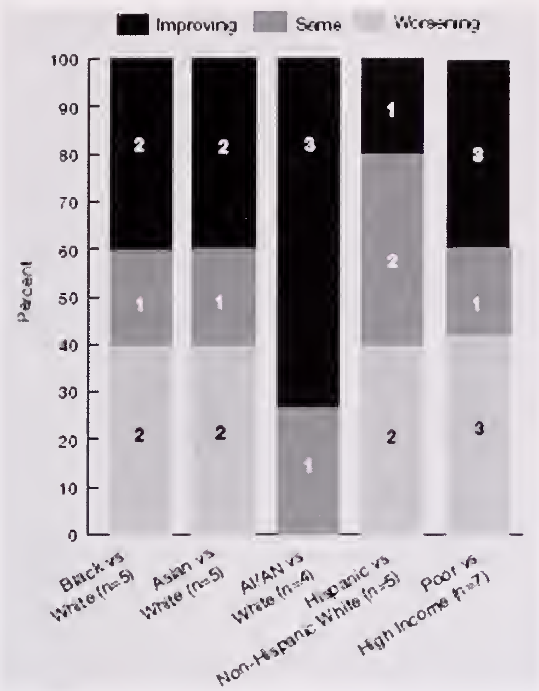


Fig 6. Changes in racial, ethnic, and income disparities in core access measures: 2000–2006. The most progress was made by AI/AN: 75% of access measures showed improvement. For blacks and Asians, only 40% of access measures improved; for Hispanics, only 20% of access measures improved. Worsening disparities in access were seen in 40% of measures for blacks, Asians, and Hispanics and in 25% of measures for AI/AN. **Abbreviations:** AI, American Indians; AN, Alaska Natives. Reprinted with permission from the Agency for Healthcare Quality and Research.[58]

quality to health care.[11–13] The report provides an illuminating breakdown of the relationships between race, ethnicity, income, education, and health insurance. Understanding these relationships is essential to efforts to eliminate racial and ethnic disparities in health care. The National Healthcare Disparities Report also monitors progress over time in eliminating racial, ethnic, and socioeconomic disparities in core measures of access and core measures of quality (see Figure 6).

Racial and Ethnic Disparities in Facilitators and Barriers to Health Care

As previously mentioned, members of racial and ethnic minorities are less likely to have health insurance.[57] Although the gap between blacks and whites in insurance coverage decreased from 1996 to 2006, the percentage of people with insurance was still lower for blacks than for whites (81.9% versus 83.3%).[58] The gap between blacks and whites in percentage of people insured all year did not change; this percentage remained higher for blacks than for whites in 2005 (15.6% versus 13.8%).[59] From 1999 to 2006, the gap between Hispanics and non-Hispanic whites in insurance coverage remained the same: in 2006, 65.0% of Hispanics versus 87.5% of non-Hispanic whites were insured. The percentage of people uninsured all year was still almost 3 times higher for Hispanics than for non-Hispanic whites in 2005 (28.5% versus 10.3%).[58]

Persons with a usual source of care (a facility at which one regularly receives care) experience improved health outcomes, reduced costs, and reduced disparities, yet over 40 million Americans do not have a specific source of ongoing care.[60–62] Compared to whites, both blacks and Hispanics are less likely to have a specific source of ongoing care[57] and less likely to have a usual primary care provider.[58] Multivariate modeling can identify the independent effects of race, ethnicity, and socioeconomic status on access to health care. For example, adjusted odds ratios from a multivariate analysis controlling for race, ethnicity, income, education, age, gender, insurance, and residence location reveal that, in comparison with whites, blacks had 0.9 times the odds and Asians had 0.7 times the odds of having a usual primary care provider, and in comparison with non-Hispanic whites, Hispanics had 0.6 times the odds of having a usual primary care provider (data from 2002–2005 for persons 18–64 years of age; see Figures 7 and 8).[58]

Racial and ethnic minorities who are able to access health services are much more likely to report poor communication with their health care provider. Many social, cultural, health literacy, linguistic, and environmental issues influence the clinical encounter and affect access to appropriate health care (see Figure 9). The success of any clinical encounter is influenced by the attitudes, expectations, and behaviors of the clinician and of the patient.[63,64]

> *The success of any clinical encounter is influenced by the attitudes, expectations, and behaviors of the clinician and of the patient.*

Competent medical decision making requires the appropriate use of 2 kinds of data: patient-specific information and the clinician's prior medical knowledge. The clinician must gather and assess the value of information about the individual patient's symptoms, medications, comorbidities, risk factors,

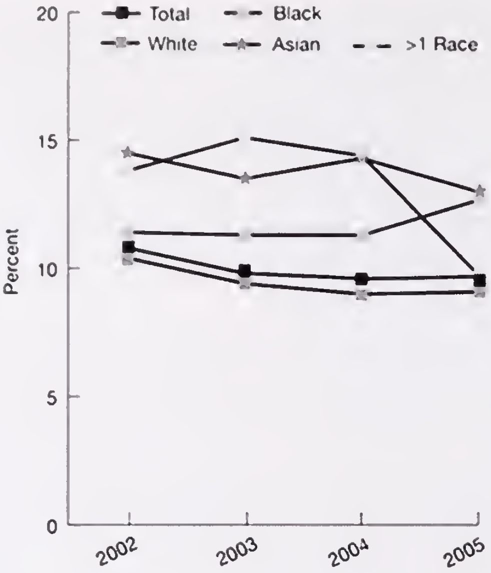


Fig 7. Ambulatory patients, 18 years old and older, who reported poor communication with health providers by race: 2002–2005. This is a composite measure that combines 4 measures of patient-provider communication: (1) providers who sometimes or never listen carefully, (2) providers who explain things clearly, (3) providers who respect what patients say, and (4) providers who spend enough time with patients. Data were insufficient for this analysis for Native Hawaiians and other Pacific Islanders, American Indians, and Alaska Natives. Adapted from *Medical Expenditure Panel Survey, 2002–2005.*[59]

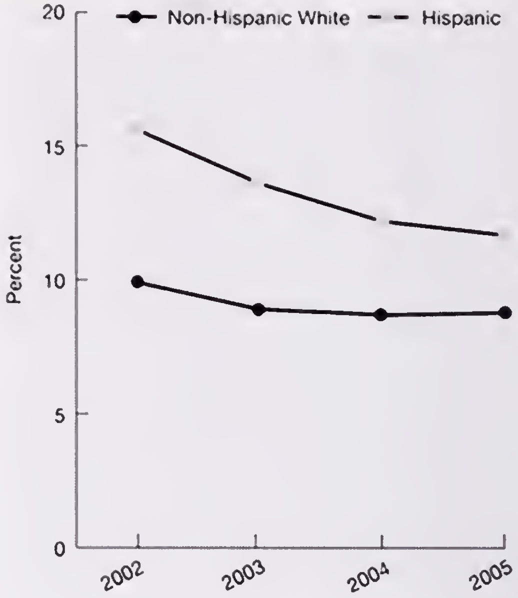


Fig 8. Ambulatory patients, 18 years old and older, who reported poor communication with health providers by ethnicity: 2002–2005. This is a composite measure that combines 4 measures of patient-provider communication: (1) providers who sometimes or never listen carefully, (2) providers who explain things clearly, (3) providers who respect what patients say, and (4) providers who spend enough time with patients. Adapted from *Medical Expenditure Panel Survey, 2002–2005.*[59]

health beliefs, and health-related behaviors. When the patient and clinician come from different backgrounds, the clinician's understanding of the information gathered from the patient may be impaired, or the clinician's interpretation may be influenced by bias; either may result in inadequate weight being placed on the data gathered from the patient.[65]

Furthermore, the provider's prior medical knowledge may not include accurate data with respect to disease presentation, prevalence, and processes in minority populations. Thus, the provider's lack of skill in acquiring patient-specific information or lack of knowledge regarding population-specific information may adversely affect the quality of diagnostic decision making for minority patients. The patient's attitudes and behavior may also adversely affect communication with the physician. Culturally based health practices, learned care-seeking behavior patterns, medical mistrust, and limited English proficiency all affect access to and utilization of health care services.[66] The resultant disparities in access to health care lead to substantial differences in health care utilization and health care outcomes, which in turn contribute to identified racial and ethnic disparities in mortality and health status.

> *The provider's lack of skill in acquiring patient-specific information or lack of knowledge regarding population-specific information may adversely affect the quality of diagnostic decision making for minority patients.*

CONCLUSION

Substantial empirical evidence documents disparate access to health and health care for members of racial and ethnic minorities. Important aspects of these disparities include differential access to health information, health promotion, prevention activities, safe housing, nutritious foods, convenient exercise spaces, freedom from ambient violence, adequate social support, and communities with social capital as well as disparate access to health insurance, a regular source of care, a usual primary care provider, and quality health care. These disparities in access to health and health care occur through a variety of

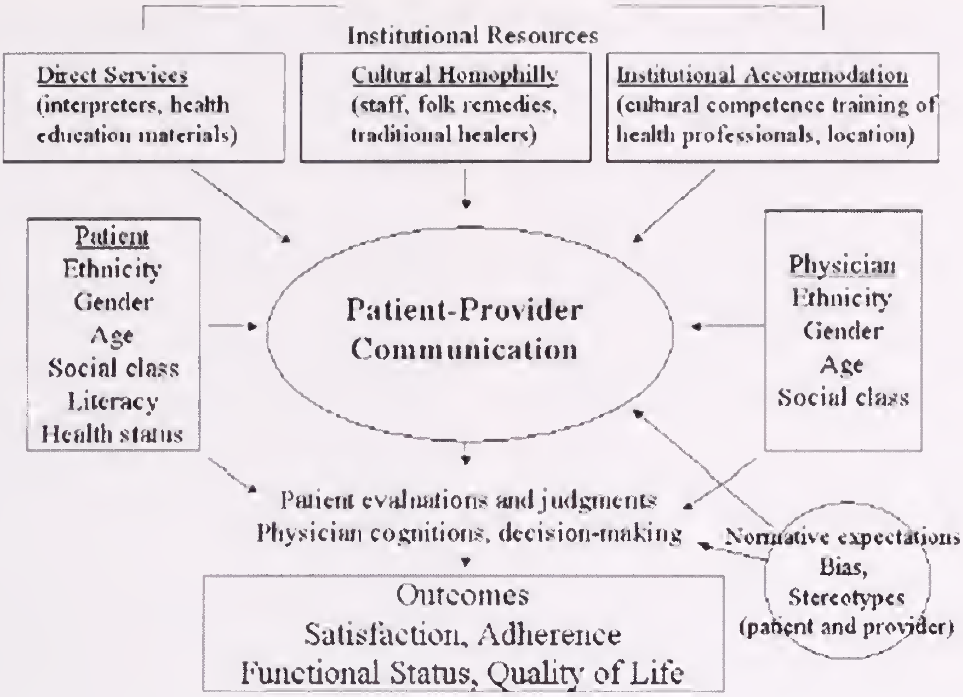

Institutional Resources
Direct Services (interpreters, health education materials)
Cultural Homophilly (staff, folk remedies, traditional healers)
Institutional Accommodation (cultural competence training of health professionals, location)
Patient
Ethnicity
Gender
Age
Social class
Literacy
Health status
Patient-Provider Communication
Physician
Ethnicity
Gender
Age
Social class
Patient evaluations and judgments
Physician cognitions, decision-making
Normative expectations
Bias,
Stereotypes
(patient and provider)
Outcomes
Satisfaction, Adherence
Functional Status, Quality of Life

Fig 9. Effects of race and ethnicity on patient-physician communication. Reprinted with permission from the National Academies Press.[64]

mechanisms, including the built environment, social environment, residential segregation, stress, racism, and discrimination.

ACKNOWLEDGMENTS

Marlaina Norris was supported by a Minority Faculty Scholar Award from the Northeast Minority Faculty Consortium and by an Empire Clinical Research Investigator Program Award from New York State.

This article is based on a presentation given by Lynne D. Richardson at the Center for Multicultural and Community Affairs Symposium on Achieving Health Equity in New York, NY on November 17, 2008.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. *Health, United States, 2008 with Chartbook.* Hyattsville, MD: National Center for Health Statistics; 2009.
2. *Health, United States, 2007 with Chartbook on Trends in the Health of Americans.* Hyattsville, MD: National Center for Health Statistics; 2007.
3. Kung HC, Hoyert DL, Xu J, Murphy SL. Deaths: final data for 2005. *Natl Vital Stat Rep* 2008; 56: 4–26.
4. *Health, United States, 2005 with Chartbook.* Hyattsville, MD: National Center for Health Statistics; 2005.
5. Williams DR. The health of U.S. racial and ethnic populations. *J Gerontol B Psychol Sci Soc Sci* 2005; 60B: 53–62.
6. Pamuk E, Makuk D, Heck K, Reuben C. *Socioeconomic Status and Health Chartbook.* Hyattsville, MD: National Center for Health Statistics; 1998.
7. Williams DR, Mohammed SA. Discrimination and racial disparities in health: evidence and needed research. *J Behav Med* 2009; 32: 20–47.
8. Williams DR, Jackson PB. Social sources of racial disparities in health. *Health Aff* 2005; 24: 325–334.
9. Health disparities: concepts, measurements, and understanding. In: Thomson GE, Mitchell F, Williams MB, eds. *Examining the Health Disparities Research Plan of the National Institutes of Health: Unfinished Business.* Washington, DC: National Academies Press; 2006: 21–33.
10. Smedley BD, Stith AY, Nelson AR, eds. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care.* Washington, DC: National Academies Press; 2003.
11. *2003 National Healthcare Disparities Report.* Rockville, MD: Agency for Healthcare Research and Quality; 2004.
12. *2007 National Healthcare Disparities Report.* Rockville, MD: Agency for Healthcare Research and Quality; 2008. AHRQ publication 08–0041.
13. *2006 National Healthcare Disparities Report.* Rockville, MD: Agency for Healthcare Research and Quality; 2007. AHRQ publication 07–0012.
14. Richardson LD, Hwang U. Access to care: a review of the emergency medicine literature. *Acad Emerg Med* 2001; 8: 1030–1036.

15. Millman M, ed. *Access to Health Care in America.* Washington, DC: National Academies Press; 1993.
16. Goodell S, Escarce JJ. *Racial and Ethnic Disparities in Access to and Quality of Health Care.* Princeton, NJ: Robert Wood Johnson Foundation; 2007. Policy brief 12.
17. Williams DR, Collins CA. Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Rep* 2001; 116: 404–415.
18. Acevedo-Garcia D, Lochner KA, Osypuk TL, Subramanian SV. Future directions in residential segregation and health research: a multilevel approach. *Am J Public Health* 2003; 93: 215–221.
19. Acevedo-Garcia D, Osypuk TL, McArdle N. Williams DR. Toward a policy-relevant analysis of geographic and racial/ethnic disparities in child health. *Health Aff* 2008; 27: 321–333.
20. CensusScope. US metro areas ranked by white/black dissimilarity index. http://www.censusscope.org/us/rank_dissimilarity_white_black.html. Accessed January 2010.
21. Polednak AP. Poverty, residential segregation, and black/white mortality rates in urban areas. *J Health Care Poor Underserved* 1993; 4: 363–373.
22. Collins CA, Williams DR. Segregation and mortality: the deadly effects of racism? *Sociol Forum* 1999; 14: 495–523.
23. Fang J, Madhavan S, Bosworth W, Alderman MH. Residential segregation and mortality in New York City. *Soc Sci Med* 1998; 47: 469–476.
24. Guest AM, Almgren G, Hussey JM. The ecology and race and socioeconomic distress: infant and working age mortality in Chicago. *Demography* 1998; 35: 25–34.
25. LaVeist TA. Linking residential segregation and infant mortality race disparity in U.S. cities. *Sociol Soc Res* 1989; 73: 90–94.
26. LaVeist TA. The political empowerment and health status of African-Americans: mapping a new territory. *Am J Sociol* 1992; 97: 1080–1095.
27. Polednak AP. Black-white differences in infant mortality in 38 standard metropolitan statistical areas. *Am J Public Health* 1991; 81: 1480–1482.
28. Acevedo-Garcia D. Zip code-level risk factors for tuberculosis: neighborhood environment and residential segregation in New Jersey, 1985–1992. *Am J Public Health* 2001; 91: 734–741.
29. Evans GW, Saegert S. Residential crowding in the context of inner city poverty. In: Wapner S, Demick J, Yamamoto T, et al., eds. *Theoretical Perspectives in Environment-Behavior Research.* New York, NY: Kluwer Academic/Plenum Publishers; 2000: 247–267.
30. Bullard RD. Urban infrastructure: social environmental, and health risks to African Americans. In: Livingston IL, ed. *Handbook of Black American Health: The Mosaic of Conditions, Issues, Policies and Prospects.* Westport, CT: Greenwood Publishing; 1994: 315–330.
31. American Lung Association. Urban air pollution and health inequities: a workshop report. *Environ Health Perspect* 2001; 109(suppl 3): 357–374.
32. Neighborhood safety and the prevalence of physical inactivity–selected states, 1996. *MMWR Morb Mortal Wkly Rep* 1999; 48: 143–146.
33. Cheadle A, Psaty BM, Curry S, et al. Community-level comparisons between the grocery store environment and individual dietary practices. *Prev Med* 1991; 20: 250–261.
34. Hacker AG, Collins R, Jacobson M. *Marketing Booze to Blacks.* Washington, DC: Center for Science in the Public Interest; 1987.
35. Rabow J, Watt R. Alcohol availability, alcohol beverage sales, and alcohol-related problems. *J Stud Alcohol* 1982; 43: 767–801.
36. Maxwell B, Jacosbon M. *Marketing Disease to Hispanics: The Selling of Alcohol, Tobacco, and Junk Foods.* Washington, DC: Center for Science in the Public Interest; 1989.
37. Mayberry RM, Price PA. Targeting blacks in cigarette billboard advertising: results from down South. *Health Values* 1993; 17: 28–35.
38. Moore AJ, Williams JD, Qualls WJ. Target marketing of tobacco and alcohol-related products to ethnic minority groups in the United States. *Ethn Dis* 1996; 6: 83–98.
39. Council of Economic Advisers for the President's Initiative on Race. *Changing America: Indicators of Social and Economic Well-Being by Race and Hispanic Origin.* Washington, DC: Government Printing Office; 1998.
40. Lillie-Blanton M, Martinez RM, Taylor AK, Robinson BG. Latina and African American women: continuing disparities in health. *Int J Health Serv* 1993; 23: 555–584.
41. Berkman LF, Glass TA. Social integration, social networks, social support, and health. In: Berkman LF, Kawachi I, eds. *Social Epidemiology.* New York, NY: Oxford University Press; 2000: 137–173.
42. Adler NE. Overview of health disparities. In: Thomson GE, Mitchell F, Williams MB, eds. *Examining the Health Disparities Research Plan of the National Institutes of Health: Unfinished Business.* Washington, DC: National Academies Press; 2006: 121–174.
43. Kawachi I, Berkman LF. Social cohesion social capital and health. In: Berkman LF, Kawachi I, eds. *Social Epidemiology.* New York, NY: Oxford University Press; 2000: 174–190.
44. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health* 2001; 55: 111–122.
45. Diez Roux AV, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. *N Engl J Med* 2001; 345: 99–106.
46. Bonilla-Silva E. Rethinking racism: toward a structural interpretation. *Am Sociol Rev* 1996; 62: 465–480.
47. Jones JM. *Prejudice and Racism.* New York, NY: McGraw-Hill; 1997.
48. Jones CP. Levels of racism: a theoretical framework and a gardener's tale. *Am J Public Health* 2000; 90: 1212–1215.
49. Jones CP, Truman BI, Elam-Evans LD, et al. Using socially assigned race to probe white advantages in health status. *Ethn Dis* 2008; 18: 496–504.
50. Williams DR, Neighbors HW, Jackson JS. Racial/ethnic discrimination and health: findings from community studies. *Am J Public Health* 2003; 93: 200–208.
51. Bierman A, Magari ES, Jette AM, et al. Assessing access as a first step toward improving the quality of care for very old adults. *J Ambul Care Manage* 1998; 121: 17–26.
52. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J Health Soc Behav* 1995; 36: 1–10.

53. Kaiser Family Foundation Commission on Medicaid and the Uninsured. *The Uninsured: A Primer*. Menlo Park, CA: Henry J. Kaiser Family Foundation; 2006.
54. Centers for Disease Control and Prevention. Self-assessed health status and selected behavioral risk factors among persons with and without health-care coverage–United States, 1994–1995. *MMWR Morb Mortal Wkly Rep* 1998; 47: 176–180.
55. Committee on the Consequences of Uninsurance. *Insuring America's Health: Principles and Recommendations*. Washington, DC: National Academies Press; 2004.
56. Hadley J. *Sicker and Poorer: The Consequences of Being Uninsured*. Washington, DC: Urban Institute; 2003.
57. Tabulations of the March current population survey files by Actuarial Research Corporation, incorporating their historical adjustments. In: *An Overview of the U.S. Health Care System: Chart Book*. Washington, DC: Centers for Medicare and Medicaid Services/Office of the Assistant Secretary for Planning and Evaluation; 2007.
58. *2008 National Healthcare Disparities Report*. Rockville, MD: Agency for Healthcare Research and Quality; 2009. AHRQ publication 09-0002.
59. *Medical Expenditure Panel Survey, 2002–2005*. Rockville, MD: Agency for Healthcare Research and Quality; 2005.
60. Starfield B, Shi L. The medical home, access to care, and insurance. *Pediatrics* 2004; 113(suppl): 1493–1498.
61. De Maeseneer JM, De Prins L, Gosset C, Heyerick J. Provider continuity in family medicine: does it make a difference for total health care costs? *Ann Fam Med* 2003; 1: 144–148.
62. US Department of Health and Human Services. *Healthy People 2010*. 2nd ed. Washington, DC: US Government Printing Office; 2000: 45.
63. Washington DL, Bowles J, Saha S, et al. Transforming clinical practice to eliminate racial-ethnic disparities in healthcare. *J Gen Intern Med* 2008; 23: 685–691.
64. Cooper LA, Roter DL. Patient-provider communication: the effect of race and ethnicity on process and outcomes of healthcare. In: Smedley BD, Stith AY, Nelson AR, eds. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington, DC: National Academies Press; 2003: 522–593.
65. Richardson LD, Irvin CB, Tamayo-Sarver JH. Racial and ethnic disparities in the clinical practice of emergency medicine. *Acad Emerg Med* 2003; 10: 1184–1188.
66. Gresenz CR, Rogowski J, Escarce JJ. Community demographics and access to health care among U.S. Hispanics. *Health Serv Res* 2009; 44(pt 1): 1542–1562.
67. Karpati A, Kerker B, Mostashari F, et al. *Health Disparities in New York City*. New York, NY: New York City Department of Health and Mental Hygiene; 2004.

Environmental Justice and the Health of Children

Philip J. Landrigan, MD, MSc,[1] Virginia A. Rauh, ScD,[2] and Maida P. Galvez, MD, MPH[1]

[1]Mount Sinai School of Medicine, New York, NY
[2]Mailman School of Public Health, Columbia University, New York, NY

OUTLINE

ENVIRONMENT AND HUMAN HEALTH
CHILDREN'S EXPOSURES TO TOXIC CHEMICALS IN THE ENVIRONMENT
WIDESPREAD FAILURE TO TEST CHEMICALS FOR TOXICITY
CHILDREN'S UNIQUE VULNERABILITY TO TOXIC CHEMICALS
EVIDENCE THAT TOXIC ENVIRONMENTAL EXPOSURES CONTRIBUTE TO MAJOR CHILDHOOD DISEASES
ENVIRONMENTAL INJUSTICE
  Asthma
  Lead Poisoning
  Obesity
CONCLUSION

ABSTRACT

Environmental injustice is the inequitable and disproportionately heavy exposure of poor, minority, and disenfranchised populations to toxic chemicals and other environmental hazards. Environmental injustice contributes to disparities in health status across populations of differing ethnicity, race, and socioeconomic status. Infants and children, because of their unique biological vulnerabilities and age-related patterns of exposure, are especially vulnerable to the health impacts of environmental injustice. These impacts are illustrated by sharp disparities across children of different racial and ethnic backgrounds in the prevalence of 3 common diseases caused in part by environmental factors: asthma, lead poisoning, and obesity. Documentation of linkages between health disparities and environmental injustice is an important step toward achieving environmental justice. ***Mt Sinai J Med 77:178–187, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: children's environmental health, environmental justice, lead poisoning, air pollution, asthma, obesity, built environment.

Address Correspondence to:

Philip J. Landrigan
Mount Sinai School of Medicine
New York, NY
Email: phil.landrigan@mssm.edu

Environmental injustice is the inequitable and disproportionately heavy exposure of poor, minority, and disenfranchised populations to toxic chemicals, contaminated air and water, unsafe workplaces, and other environmental hazards.[1–3] The concept of environmental injustice was first developed in the 1980s in studies of hazardous waste sites in the Southeastern United States.[1] These studies found that waste sites in the Southeast are located disproportionately in poor counties inhabited largely by African Americans, Native Americans, and other marginalized populations. A similar distribution of hazardous waste sites was subsequently documented in New England.[4] The concept of environmental injustice has been further elaborated in studies examining ethnic disparities in exposures to automotive exhaust and ambient air pollution[5–7]; in studies in New York City documenting that virtually all diesel bus depots, places at which buses may idle for hours while emitting pollutants, are located in minority, mostly disadvantaged

> *Environmental injustice is the inequitable and disproportionately heavy exposure of poor, minority, and disenfranchised populations to toxic chemicals, contaminated air and water, unsafe workplaces, and other environmental hazards.*

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20173

© 2010 Mount Sinai School of Medicine

neighborhoods[8]; in studies examining disparities in housing quality[9]; and in studies of residential proximity to polluting industrial facilities.[10] Environmental injustice has been well documented in occupational settings.[11] It has served as an operational concept to guide pollution prevention programs.[12,13]

Environmental injustice is highly correlated with other factors that link poverty to poor health, including inadequate access to medical and preventive care, lack of safe play spaces for children, lack of access to healthful foods, absence of good jobs, crime, and violence.[2] Environmental injustice contributes to disparities in health status across populations of different ethnic, racial, and socioeconomic backgrounds, such as differences in the incidence and prevalence of asthma,[5–7] obesity,[14,15] diabetes,[16] lung cancer,[5] and a range of mental health and developmental problems.[17–19]

Children are especially vulnerable to the health impacts of environmental injustice. Children's unique, age-related patterns of exposure and their developmentally defined windows of susceptibility[20] magnify the impacts of environmental injustice. Through case studies of 3 diseases caused entirely or in part by environmental exposures–asthma, lead poisoning, and obesity–this report examines the consequences of toxic environmental exposures and environmental injustice for the health of children.

Children are especially vulnerable to the health impacts of environmental injustice. Children's unique, age-related patterns of exposure and their developmentally defined windows of susceptibility magnify the impacts of environmental injustice.

ENVIRONMENT AND HUMAN HEALTH

Hazardous exposures in the environment are potent causes of disease, disability, and death in persons of all ages and especially in infants and children.

The power of the environment to influence patterns of disease and death is illustrated by the extraordinary changes in morbidity and mortality that have occurred over the past century in industrially developed countries, changes unprecedented in human history. Life expectancy at birth has increased more than 50%.[21] Infant mortality has declined more than 90%. The ancient infectious diseases–smallpox, cholera, yellow fever, polio, measles, and bubonic plague–are no longer the dominant causes of disease and death.[22]

These changes occurred in parallel with large-scale environmental changes–the delivery of safe drinking water; the provision of sufficient, wholesome food; the removal of sewage; the control of insect vectors; and the construction of decent housing–and were largely driven by these changes. It is noteworthy that the decline in mortality that marked the start of the epidemiological transition began in the United States in the 1860s, soon after the construction of major urban water systems and nearly 80 years before the discovery of penicillin and more than a century before the first organ transplant (see Figure 1).

Today, in the aftermath of the epidemiological revolution, the principal diseases of American children are a group of chronic diseases termed the new pediatric morbidity.[23] These diseases are the major causes of illness and death in American children today:

- Asthma, which has more than doubled in frequency since 1980 and become the leading cause of pediatric hospitalization and school absenteeism of American children. Rates of asthma have risen especially rapidly among poor children of color residing in inner-city communities.[24]
- Birth defects, which are now the leading cause of infant death. Certain birth defects of the male reproductive organs, such as hypospadias, have doubled in frequency.[25,26]
- Neurodevelopmental disorders, such as dyslexia, mental retardation, attention deficit/hyperactivity disorder, and autism. These conditions affect 5% to 10% of the 4 million babies born each year in the United States. The reported frequency is sharply increasing.[27]
- Leukemia and brain cancer in children, which have increased in reported incidence since the 1970s, despite declining mortality.[28,29] Cancer is now the second leading cause of death in American children and is surpassed only by traumatic injuries.
- Preterm birth, which has increased in incidence by 27% since 1981.
- Obesity, which has trebled in prevalence, and its result, type 2 diabetes. In 2005, 41% of 5-year-old children entering kindergarten in the 5 boroughs of New York City were found to be overweight or obese. Obesity is especially prevalent in African American and Latino children.[30]

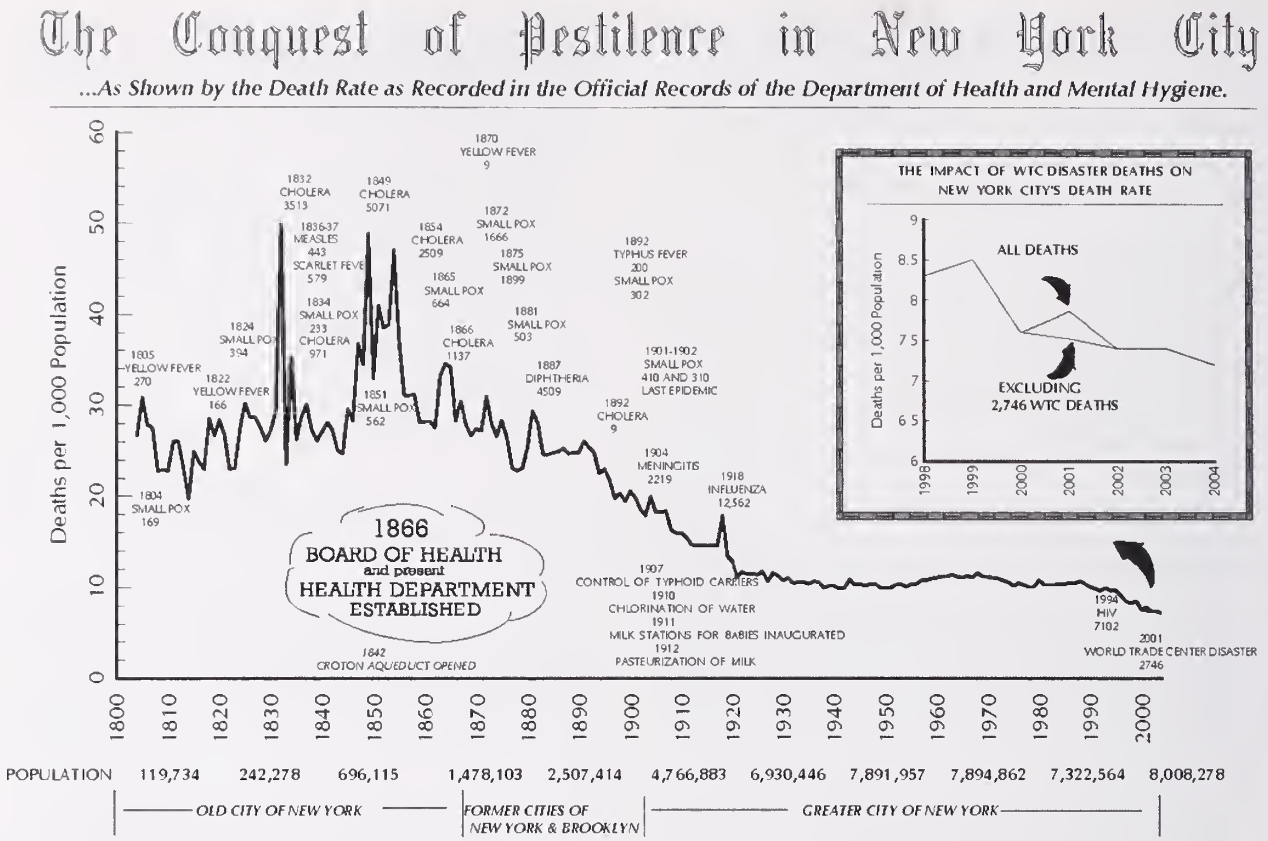


Fig 1. Conquest of pestilence in New York City. **Abbreviations:** HIV, human immunodeficiency virus; WTC, World Trade Center.

CHILDREN'S EXPOSURES TO TOXIC CHEMICALS IN THE ENVIRONMENT

> *Children today are at risk of exposure to more than 80,000 synthetic chemicals. Most of these chemicals are newly invented, and nearly all of them did not exist 50 years ago.*

Children's environments have changed profoundly in the past century. Children today are at risk of exposure to more than 80,000 synthetic chemicals. Most of these chemicals are newly invented, and nearly all of them did not exist 50 years ago. They include plastics, pesticides, motor fuels, building materials, antibiotics, chemotherapeutic agents, flame retardants, and synthetic hormones.[31] Children are especially at risk of exposure to the 3000 synthetic chemicals that are produced in quantities of 1 million pounds or more per year. These high production volume (HPV) chemicals are the synthetic materials used most extensively in industry and consumer products and most widely dispersed in the environment: in air, food, water, homes, schools, and communities. Measurable levels of several hundred HPV chemicals have been documented in Centers for Disease Control and Prevention surveys in the bodies of most Americans.[32] Measurable levels of HPV chemicals have been documented also in the breast milk of nursing mothers and in the cord blood of newborn infants.[33]

WIDESPREAD FAILURE TO TEST CHEMICALS FOR TOXICITY

A high proportion of the most widely used synthetic chemicals have never been tested for their possible toxicity.[34] Information on potential toxicity is publicly available for only about two-thirds of the 3000 HPV chemicals. Information on possible developmental toxicity or the potential capacity to cause injury to infants and children is especially lacking. This information is available for less than one-third of HPV chemicals.[34]

Failure to test chemicals for toxicity represents a serious lapse of responsible stewardship. It puts children at risk on a daily basis to exposure to chemicals whose hazardous potential is virtually unknown. It reflects a combination of industry's unwillingness to take responsibility for the products that it produces coupled with a failure of the government to regulate the use of these products.

Governmental failure in this area flows from the failure of the Toxic Substances Control Act (TSCA).[35] TSCA was intended at the time of its passage in 1976 to be pioneering legislation that would require premarket evaluation of all new chemicals for potential toxicity and also require retroactive testing of tens of thousands of industrial chemicals that were already in commerce. In fact, however, TSCA has been a dismal failure. A particularly egregious example of this failure was a decision made soon after the passage of TSCA to grandfather in, with no toxicity testing, 62,000 chemicals that were already on the market. These chemicals were simply presumed to be safe and allowed to remain in commerce. The office within the US Environmental Protection Agency responsible for enforcing TSCA has been chronically underfunded, understaffed, and overwhelmed by the sheer number of new chemicals and technologies that come before it. By default, emerging chemicals and new products are presumed by Environmental Protection Agency regulators to be safe unless there is overwhelming evidence of their potential to cause harm.

CHILDREN'S UNIQUE VULNERABILITY TO TOXIC CHEMICALS

Children are now understood to be fundamentally more vulnerable than adults to toxic chemicals in the environment.[20] Four differences between children and adults contribute to children's increased susceptibility:

- Children have disproportionately heavier exposures to chemicals in comparison with adults. This reflects children's greater consumption per pound of body weight of food, water, and air. Thus, a child in the first year of life drinks 7 times more water per pound per day than an adult. These dietary exposures are further magnified by children's unique behaviors: their play close to the floor and their oral-exploratory behavior both permit all too easy access to toxic materials in rugs, dust, and soil.
- Children's metabolic pathways, especially in the first months after birth, are immature. In many instances, children are less able than adults to break down and excrete toxic compounds. Thus, organophosphate pesticides linger in the bloodstream of a child for 36 hours, whereas most adults are able to clear and excrete these dangerous materials in 6 hours and thus minimize damage.
- Children are undergoing rapid growth and development. Early development creates windows of great vulnerability. Witness, for example, the uniquely tragic consequences of exposures in early life to substances such as thalidomide, diethylstilbestrol, and methyl mercury.
- Because children have more years of future life than most adults, they have more time to develop chronic diseases that may be initiated by early exposures.

EVIDENCE THAT TOXIC ENVIRONMENTAL EXPOSURES CONTRIBUTE TO MAJOR CHILDHOOD DISEASES

Evidence has been building for 100 years that toxic exposures in the environment can cause and aggravate disease in children.[36] This body of evidence has become especially compelling in the past 2 decades:

- Rates of asthma are increased in children exposed to second-hand cigarette smoke.[37]
- Rates of asthma are increased in children exposed to particulate air pollution.[38–40]
- Risk of sudden infant death syndrome is increased in babies exposed to particulate air pollution.[41]
- Neurodevelopmental impairment with a reduced intelligence quotient (IQ), shortening of attention span, and disruption of behavior is increased in children exposed to lead.[42–46]
- Neurodevelopmental impairment with a reduction of IQ is increased in children exposed to polychlorinated biphenyls.[47]
- Neurodevelopmental impairment with a reduction of IQ and shortening of attention span is increased in children exposed prenatally to methyl mercury.[48]
- Lower birth weight, smaller head circumference at birth, and subsequent developmental delays are increased in children exposed prenatally to the organophosphate pesticide chlorpyrifos.[49,50]
- Neurodevelopmental impairment with a reduction of IQ and disruption of behavior is increased in children exposed prenatally to ethyl alcohol; this is called fetal alcohol syndrome.[51]

- There is an increased risk of preterm birth associated with prenatal exposure to tobacco smoke and dichlorodiphenyltrichloroethane.[52]

Many of these epidemiological findings have been translated into evidence-based programs of prevention. These prevention programs have yielded enormous gains for children's health. Examples include the following:

- The finding that alcohol and tobacco exposure during pregnancy can cause brain damage and growth restriction in infants has been successfully translated into public health intervention programs that have reduced alcohol consumption and tobacco use during pregnancy. These programs have led to the prevention of fetal alcohol syndrome and intrauterine growth retardation.[53,54]
- The finding that X-ray exposure during pregnancy increases the risk of childhood cancer has triggered minimization during pregnancy of diagnostic X-rays. This change in medical practice has prevented thousands of cases of childhood leukemia.[55]
- Removal of lead from gasoline, following the discovery that low-dose exposures to lead could cause infant brain damage with a loss of intelligence and disruption of behavior, has resulted in a 90% reduction in childhood blood lead levels and lead poisoning.[43]
- A reduction in children's exposure to arsenic in well water has brought about a reduction in the risk of skin and liver cancer in exposed children.[56]

ENVIRONMENTAL INJUSTICE

> *Well-studied examples of environmental injustice in the United States include the disproportionate exposure of poor and minority communities to hazardous waste sites, lead paint, air pollution, substandard housing, dangerous jobs, and polluting industries.*

Well-studied examples of environmental injustice in the United States include the disproportionate exposure of poor and minority communities to hazardous waste sites, lead paint, air pollution, substandard housing, dangerous jobs, and polluting industries. Such environmental risk factors are disproportionately concentrated in poor communities and in communities inhabited by people of color.[1–13]

Disproportionate exposures of persons in poor communities and in communities inhabited by people of color to toxic environmental hazards appear to contribute significantly to well-documented disparities in health. Environmental injustice is linked to increased risks of many diseases. These linkages are especially evident in children because of children's unique vulnerabilities to environmental hazards.[20] Three examples are now discussed.

Asthma

Asthma is the most common chronic disease of American children and is the leading cause of pediatric emergency department visits, pediatric hospitalizations, and school absenteeism. Well-described environmental risk factors for asthma include ambient air pollution from industrial and vehicular sources, indoor air pollution, secondhand cigarette smoke, mold, mildew, and cockroach droppings.[57]

The incidence, prevalence, and hospitalization rates for asthma in the United States are all disproportionately high in poor communities.[58] The highest rates are seen among poor minority children living in inner-city communities. For example, the prevalence of childhood asthma in certain low-income, minority neighborhoods in New York is as high as 23%, which is approximately 4 times the national average. Current US asthma prevalence is

- Higher in blacks (9.2%) than in whites (6.9%).
- Higher in persons of Puerto Rican descent (14.5%) than in those of Mexican descent (3.9%).
- Higher in those below the federal poverty level (10.3%) than in those at or above the federal poverty level (6.4%–7.9%).[57,59,60]

Disproportionate exposures to ambient air pollutants in poor communities are a clear example of environmental injustice associated with elevated rates of asthma. Among the ambient air pollutants found in especially high concentrations in poor, minority neighborhoods are fine particulates, ozone, and oxides of nitrogen. Bus depots, in which diesel buses were historically permitted to idle for hours on end, are an important source of ambient air pollutants in poor neighborhoods. Six of the 7 bus depots in all of Manhattan are located in either Harlem or East Harlem.

Disproportionate exposure to indoor air pollution in poor neighborhoods further increases the risk of asthma.[61,62] Among the indoor air pollutants documented to be disproportionately present in poor communities are environmental (second-hand) tobacco smoke, mold and mildew, animal dander, and cockroach droppings.

Public housing may be a particular risk factor for asthma and a further example of environmental injustice. In New York City, for example, within communities with high asthma rates, asthmatics were 5 times more likely than nonasthmatics to live in public housing. Also, short-term housing used by transient individuals is likely to be characterized by poorer quality management and upkeep, with conditions that can trigger asthma symptoms in those without permanent homes.[63,64]

Lead Poisoning

Lead is one of the best known and most extensively studied of the toxic chemicals in the environment.[42]

The most important environmental source of lead exposure for children in the United States is lead-based paint in housing built before 1978 and the lead-contaminated dust that arises from the abrasion, flaking, and chipping of lead-based paint. The major route of children's exposure to lead from paint is the ingestion of lead-contaminated dust. Children between the ages of 1 and 6 years are at the highest risk of lead exposure because the oral-exploratory, hand-to-mouth behavior that is so normal in this age group facilitates the transfer of lead dust from the environment into children's bodies.

Lead is now understood to be toxic to children at every level of exposure. High-level exposure causes acute poisoning with coma and convulsions. Lower level exposure, too low to produce symptoms, still causes injury to the brain and other organ systems.[42] Diminished intelligence, shortening of attention span, reading problems, attention deficit/hyperactivity disorder, school failure, delinquency, and criminal behavior are the consequences of exposure to lead.[44,45] These effects are lifelong, and they cannot be cured by any current medical treatment. There is no evidence of a threshold level below which lead does not cause these harmful effects on children's health.[46]

Lead poisoning is disproportionately concentrated in poor minority communities in the United States (see Figure 2). This reflects the fact that older (pre-1978) housing units that are in poor repair are disproportionately concentrated in these neighborhoods. The resulting environmental injustice is documented in the sharp disparities observed in the distribution of elevated blood lead levels in American children by race and income. Thus, 4.6% of African American children are estimated to have blood lead concentrations above 25 μg/dL versus 1.2% of white children.[43] In New York City, lead poisoning is found almost exclusively among African American and Hispanic children.

Obesity

Obesity is the single childhood condition in the United States that is rapidly worsening.[65–67] It is estimated that 17% of children 6 to 11 years old are obese versus 4% during 1971–1974.

Sharp disparities exist by race, income, and ethnicity in the prevalence of childhood obesity. It is well documented that minority children, specifically Mexican Americans and non-Hispanic blacks, have a higher prevalence of obesity than non-Hispanic white populations.[68] Prevalence rates of childhood obesity in New York City public elementary school students further illustrate these national trends, with Latinos (31%) and African Americans (23%) at disproportionate risk in comparison with their white counterparts (16%).[69] These very same minority populations are at higher risk for obesity-related sequelae, including cardiovascular disease, stroke, type 2 diabetes, and metabolic syndrome.[68,69–75]

Ecological systems theory highlights the importance of considering the role of the ecological niche in order to understand health outcomes.[76] When ecological systems theory is being applied to childhood obesity, Davison and Birch[76] recommend the consideration of contextual factors within children's families and the school, community, and society at large that specifically influence children's dietary and physical activity behaviors and hence risk for obesity. These factors, often called in the literature the built environment, may include crowded streets, a lack of outdoor play spaces, easy access to fast foods and junk foods, limited access to fresh, healthy foods, and substandard housing.[77] Extensive research has documented that each of these risk factors is disproportionately common in poor neighborhoods,[78,79] the very same communities already at increased risk for obesity.

The built environment research has taken a multilevel approach to tackling the obesity epidemic by searching for evidence-based interventions and health policies that promote healthy communities. These novel approaches have focused on diverse environmental risk factors and include reducing access to unhealthy snacks by eliminating soda and snack machines from schools nationwide,[80] banning transfats from restaurants and requiring fast food restaurants to post calorie counts,[81] increasing access to green markets,[82] providing

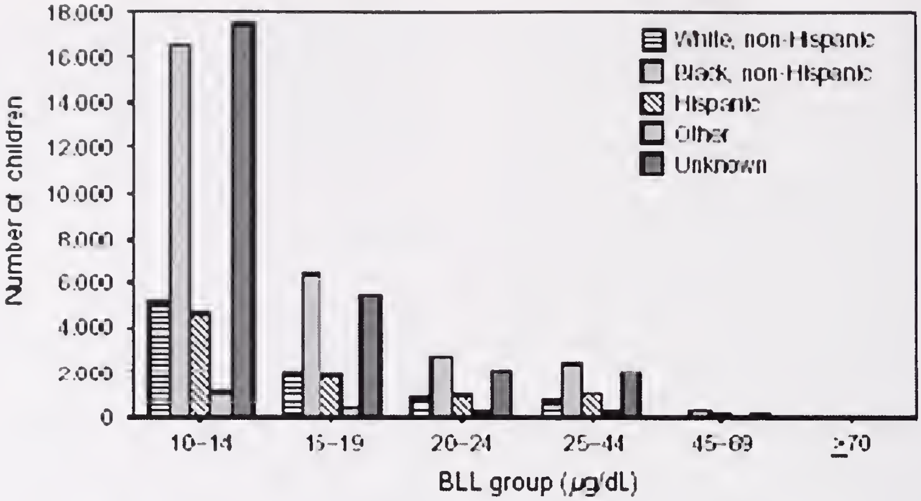


Fig 2. Number of children with confirmed BLLs by racial or ethnic groups and program-relevant BLL groups: selected US states, 2001. **Abbreviation:** BLL, blood lead level.

walking and bicycling trails,[83] adopting bike share programs in urban areas,[84] altering parks and other recreational spaces to optimize activity levels,[85] and developing walkable communities.[86] Information gained from further interdisciplinary, longitudinal cohort studies of children can inform multifaceted obesity interventions such as food policy and urban planning initiatives that promote public health.

Recent research is also examining microenvironmental factors and specifically exposure to chemicals with the potential to mimic hormones, which are commonly called endocrine disruptors.[87–89] Reports of endocrine disruptors such as phthalates and bisphenol A have been widespread in the media because of concerns that universal exposures in the US population, with the highest levels typically found in children, may be associated with health effects, including obesity and diabetes.[90,91] Although further human studies are needed, this demonstrates another potentially important area of environmental intervention with respect to tackling the epidemic of childhood obesity.

CONCLUSION

Disease of environmental origin is preventable. Disease caused by environmental degradation and environmental injustice is the result of human activity and is therefore preventable through modification or cessation of the activity that damages or pollutes the environment.

To begin to address environmental injustice and its consequences for human health, we need solid evidence-based research on the long-term human health effects of exposure to a range of physical, chemical, and social exposures beginning before conception and continuing throughout childhood into adulthood.[92–94] Prospective epidemiological studies assessing exposure-health associations, including biomarkers of exposure and finely tuned neuropsychological and behavioral endpoints, are desperately needed. These studies must be conducted within and across populations to document inequalities in exposure as well as susceptibility to the effects of toxic exposures. They need to be designed and implemented with strong community participation at every stage.

In addition, we need to recognize that inequalities in exposures arise at both the community and individual levels, so efforts to redress these disparities cannot be limited to behavioral interventions. Broad societal efforts and a reordering of priorities are needed to redress social and structural conditions that result in unacceptably high levels of toxic exposure for whole populations.[94] Protection of the vulnerable must become a moral and ethical beacon that guides decision making at all levels of government and business. An example of such enlightened decision making is to be found in the Food Quality Protection Act of 1996.[95] This is the major federal pesticide law in the United States, and it is the only federal environmental statute that makes explicit provision for the protection of children's health in the setting of pesticide standards.

As a result of environmental injustice, too many American children live their lives burdened by biological and social conditions that severely limit their potential for success. The effects on children are the cruelest consequences of environmental injustice.

Beyond childhood, the disproportionate exposure in early life of children in poor communities to environmental hazards sets the stage for a lifetime of suboptimal health and diminished achievement and thus helps to perpetuate the intergenerational cycle of disenfranchisement and poverty.[96]

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Bullard RD, Wright BH. Environmental justice for all: community perspectives on health and research needs. *Toxicol Ind Health* 1993; 9: 821–841.
2. Brulle RJ, Pellow DN. Environmental justice: human health and environmental inequalities. *Annu Rev Public Health* 2006; 27: 103–124.
3. Birnbaum LS, Zenick H, Brance CM. Environmental justice: a continuing commitment to an evolving concept. *Am J Public Health* 2009; 99(suppl 3): S487–S489.
4. Faber DR, Krieg EJ. Unequal exposure to ecological hazards: environmental injustices in the Commonwealth of Massachusetts. *Environ Health Perspect* 2002; 110(suppl 2): 277–288.
5. McEntee JC, Ogneva-Himmelberger Y. Diesel particulate matter, lung cancer, and asthma incidences along major traffic corridors in MA, USA: a GIS analysis. *Health Place* 2008; 14: 817–828.
6. Chakraborty J, Zandbergen PA. Children at risk: measuring racial/ethnic disparities in potential exposure to air pollution at school and home. *J Epidemiol Community Health* 2007; 61: 1074–1079.
7. Maantay J. Asthma and air pollution in the Bronx: methodological and data considerations in using GIS for environmental justice and health research. *Health Place* 2007; 13: 32–56.
8. Vásquez VB, Minkler M, Shepard P. Promoting environmental health policy through community based participatory research: a case study from Harlem, New York. *J Urban Health* 2006; 83: 101–110.
9. Rauh VA, Landrigan PJ, Claudio L. Housing and health: intersection of poverty and environmental exposures. *Ann N Y Acad Sci* 2008; 1136: 276–288.
10. Mohai P, Lantz PM, Morenoff J, et al. Racial and socioeconomic disparities in residential proximity to polluting industrial facilities: evidence from the Americans' Changing Lives Study. *Am J Public Health* 2009; 99(suppl 3): S649–S656.
11. Farquhar SA, Goff NM, Shadbeh N, et al. Occupational health and safety status of indigenous and Latino farmworkers in Oregon. *J Agric Saf Health* 2009; 15: 89–102.
12. Sexton K, Olden K, Johnson BL. "Environmental justice": the central role of research in establishing a credible scientific foundation for informed decision making. *Toxicol Ind Health* 1993; 9: 685–727.
13. Shendell DG, Rawling MM, Foster C, et al. The outdoor air quality flag program in central California: a school-based educational intervention to potentially help reduce children's exposure to environmental asthma triggers. *J Environ Health* 2007; 70: 28–31.
14. Sturm R, Datar A. Body mass index in elementary school children, metropolitan area food prices, and food outlet density. *Public Health* 2005; 119: 1059–1068.
15. Sturm R. Disparities in the food environment surrounding US middle and high schools. *Public Health* 2008; 122: 681–690.
16. Horowitz CR, Colson KA, Hebert PL, Lancaster K. Barriers to buying healthy foods for people with diabetes: evidence of environmental disparities. *Am J Public Health* 2004; 94: 1549–1554.
17. Xue Y, Leventhal T, Brooks-Gunn J, Earls FJ. Neighborhood residence and mental health problems of 5- to 11-year-olds. *Arch Gen Psychiatry* 2005; 62: 554–563.
18. Spencer N. Social, economic, and political determinants of child health. *Pediatrics* 2003; 112(pt 2): 704–706.
19. Schneiders J, Drukker M, van der Ende J, et al. Neighbourhood socioeconomic disadvantage and behavioural problems from late childhood into early adolescence. *J Epidemiol Community Health* 2003; 57: 699–703.
20. National Academy of Sciences. *Pesticides in the Diets of Infants and Children*. Washington, DC: National Academy Press; 1993.
21. Hoyert DL, Matthews TJ, Meacker F, et al. Annual summary of vital statistics: 2004. *Pediatrics* 2006; 117: 168–183.
22. Omran AR. The epidemiologic transition: a theory of the epidemiology of population change. *Milbank Mem Fund Q* 1971; 49: 509–538.
23. Haggerty R, Rothman J. *Child Health and the Community*. New York, NY: John Wiley & Sons; 1975.
24. Akinbami LJ, Schoendorf KC. Trends in childhood asthma: prevalence, health care utilization and mortality. *Pediatrics* 2002; 110(pt 1): 315–322.
25. Centers for Disease Control and Prevention. Improved national prevalence estimates for 18 selected major birth defects–United States, 1999–2001. *MMWR Morb Mortal Wkly Rep* 2006; 54: 1301–1305.
26. Paulozzi LJ, Erickson JD, Jackson RJ. Hypospadias trends in two US surveillance systems. *Pediatrics* 1997; 100: 831–834.
27. Boyle CA, Cordero JF. Birth defects and disabilities: a public health issue for the 21st century. *Am J Public Health* 2005; 95: 1884–1886.
28. National Cancer Institute. SEER database. http://seer.cancer.gov. Accessed January 2010.
29. Devesa SS, Blot WJ, Stone BJ, et al. Recent cancer trends in the United States. *J Natl Cancer Inst* 1995; 87: 175–182.
30. Ogden CL, Carroll MD, Curtin LR, et al. Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA* 2006; 295: 1549–1555.
31. *Chemicals in Commerce Information System (Chemical Update System Database)*. Washington, DC: US Environmental Protection Agency; 1998.
32. Centers for Disease Control and Prevention. *Third National Report on Human Exposure to Environmental Chemicals*. Atlanta, GA: National Center for Environmental Health; 2005. http://www.cdc.gov/exposurereport/. Accessed January 2010.

33. Houlihan J, Kropp T, Wiles R, et al. Body burden: the pollution in newborns. http://www.ewg.org/reports/bodyburden2/execsumm.php. Published July 14, 2005. Accessed January 2010.
34. Goldman LR. Chemicals and children's environment: what we don't know about risks. *Environ Health Perspect* 1998; 106(suppl 3): 875–880.
35. US Congress. Toxic Substances Control Act [as amended through P.L. 107–377, December 31, 2002]. http://epw.senate.gov/tsca.pdf. Accessed January 2010.
36. Miller RW. How environmental hazards in childhood have been discovered: carcinogens, teratogens, neurotoxicants, and others. *Pediatrics* 2004; 113(suppl): 945–951.
37. Salam MT, Li YF, Langholz B, Gilliland FD. Early-life environmental risk factors for asthma: findings from the Children's Health Study. *Lancet* 2004; 363: 119–125.
38. Gauderman WJ, Avol E, Gilliland F, et al. The effect of air pollution on lung development from 10 to 18 years of age. *N Engl J Med* 2004; 351: 1057–1067.
39. Friedman MS, Powell KE, Hutwagner L, et al. Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma. *JAMA* 2001; 285: 897–905.
40. Suh HH, Bahadori T, Vallarino J, Spengler JD. Criteria air pollutants and toxic air pollutants. *Environ Health Perspect* 2000; 108(suppl 4): 625–633.
41. Woodruff T, Grillo J, Schoendorf C. The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. *Environ Health Perspect* 1997; 105: 608–612.
42. Agency for Toxic Substances and Disease Registry. *Toxicological Profile for Lead*. US Department of Health and Human Services: Atlanta, GA; 1999.
43. Centers for Disease Control and Prevention. Blood lead levels in young children–United States and selected states, 1996–1999. *MMWR Morb Mortal Wkly Rep* 2000; 49: 1133–1137.
44. Needleman HL, Gunnoe C, Leviton A, et al. Deficits in psychologic and classroom performance of children with elevated dentine lead levels. *N Engl J Med* 1979; 300: 689–695.
45. Bellinger DC, Stiles KM, Needleman HL. Low-level lead exposure, intelligence and academic achievement: a long-term follow-up study. *Pediatrics* 1992; 90: 855–861.
46. Canfield RL, Henderson CR Jr, Cory-Slechta DA, et al. Intellectual impairment in children with blood lead concentrations below 10 microg per deciliter. *N Engl J Med* 2003; 348: 1517–1526.
47. Jacobson JL, Jacobson SW. Intellectual impairment in children exposed to polychlorinated biphenyls in utero. *N Engl J Med* 1996; 335: 783–789.
48. Grandjean P, Weihe P, White RF, et al. Cognitive deficit in 7-year old children with prenatal exposure to methylmercury. *Neurotoxicol Teratol* 1997; 19: 417–428.
49. Whyatt RM, Rauh V, Barr DB, et al. Prenatal insecticide exposures and birth weight and length among an urban minority cohort. *Environ Health Perspect* 2004; 112: 1125–1132.
50. Rauh V, Garfinkel R, Perera FP, et al. Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first three years of life among inner-city children. *Pediatrics* 2006; 118: e1845–e1859.
51. Barr HM, Streissguth AP. Identifying maternal self-reported alcohol use associated with fetal alcohol spectrum disorders. *Alcohol Clin Exp Res* 2001; 25: 283–287.
52. Longnecker MP, Klebanoff MA, Zhou H, Brock JW. Association between maternal serum concentration of the DDT metabolite DDE and preterm and small-for-gestational-age babies at birth. *Lancet* 2001; 358: 110–114.
53. Olsen H, Burgess D, Streissguth A. Fetal alcohol syndrome and fetal alcohol effects: a lifespan view, with implications for intervention. *Zero Three/Natl Cent Clin Infant Programs* 1992; 13: 24–29.
54. Lumley J, Oliver SS, Chamberlain C, Oakley L. Interventions for promoting smoking cessation during pregnancy. *Cochrane Database Syst Rev* 2004; 3: CD001055.
55. Stewart AM. Leukemia and other neoplasms in childhood following radiation exposure in utero–a general survey of present knowledge. *Br J Radiol* 1968; 41: 718–719.
56. Graziano JH, van Geen A. Reducing arsenic exposure from drinking water: different settings call for different approaches. *Environ Health Perspect* 2005; 113: A360–A361.
57. Moorman JE, Rudd RA, Johnson CA. National surveillance for asthma–United States, 1980–2004. *MMWR Surveill Summ* 2007; 56: 1–54.
58. Claudio L, Tulton L, Doucette J, Landrigan PJ. Socioeconomic factors and asthma hospitalization rates in New York City. *J Asthma* 1999; 36: 343–350.
59. Corburn J, Osleeb J, Porter M. Urban asthma and the neighborhood environment in New York City. *Health Place* 2006; 12: 167–179.
60. Gold DR, Wright R. Population disparities in asthma. *Ann Rev Public Health* 2005; 26: 89–113.
61. Brugge DJ, Valiarino J, Ascolillo L. Comparison of multiple environmental factors for asthmatic children in public housing. *Indoor Air* 2003; 13: 18–27.
62. Clougherty JE, Levy JI, Kubzansky LD, et al. Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology. *Environ Health Perspect* 2007; 115: 1140–1146.
63. Levy J, Brugge D, Peters JL. A community-based participatory research study of multifaceted in-home environmental interventions for pediatric asthmatics in public housing. *Soc Sci Med* 2005; 63: 2191–2203.
64. Peters JL, Levy JI, Rogers CA. Determinants of allergen concentrations in apartments of asthmatic children living in public housing. *J Urban Health* 2007; 84: 185–197.
65. US Department of Health and Human Services. Healthy People 2020. http://www.healthypeople.gov/HP2020. Accessed January 2010.
66. Ogden CL, Flegal KM, Carroll MD, Johnson CL. Prevalence and trends in overweight among US children and adolescents, 1999–2000. *JAMA* 2002; 288: 1728–1732.
67. Ogden CL, Carroll MD, Flegal KM. High body mass index for age among US children and adolescents 2003–2006. *JAMA* 2008; 299: 2401–2405.
68. Ogden CL, Carroll MD, Curtin LR, et al. Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA* 2006; 295: 1549–1555.

69. Thorpe LE, List DG, Marx T, et al. Childhood obesity in New York City elementary school students. *Am J Public Health* 2004; 94: 1496–1500.
70. Shaw J. Epidemiology of childhood type 2 diabetes and obesity. *Pediatr Diabetes* 2007; 8(suppl 9): 7–15.
71. James WP. The epidemiology of obesity: the size of the problem. *J Intern Med* 2008; 263: 336–352.
72. Lee JM. Why young adults hold the key to assessing the obesity epidemic in children. *Arch Pediatr Adolesc Med* 2008; 162: 682–687.
73. Morrison JA, Ford ES, Steinberger J. The pediatric metabolic syndrome. *Minerva Med* 2008; 99: 269–287.
74. Zimmet P, Alberti G, Kaufman F, et al. The metabolic syndrome in children and adolescents. *Lancet* 2007; 369: 2059–2061.
75. Zimmet P, Alberti KG, Kaufman F, et al. The metabolic syndrome in children and adolescents–an IDF consensus report. *Pediatr Diabetes* 2007; 8: 299–306.
76. Davison KK, Birch LL. Childhood overweight: a contextual model and recommendations for future research. *Obes Rev* 2001; 2: 159–171.
77. Jackson RJ. The impact of the built environment on health: an emerging field. *Am J Public Health* 2003; 93: 1382–1384.
78. Galvez MP, Frieden TR, Landrigan PJ. Obesity in the 21st century. *Environ Health Perspect* 2003; 111: A684–A685.
79. Giles-Corti B, Macintyre S, Clarkson JP, et al. Environmental and lifestyle factors associated with overweight and obesity in Perth, Australia. *Am J Health Promot* 2003; 18: 93–102.
80. Winterfeld A. Nutrition rules: making healthy food choices available to school kids is a priority for many lawmakers. *State Legis* 2008; 34: 22–24.
81. Frieden TR, Bassett MT, Thorpe LE, Farley TA. Public health in New York City, 2002–2007: confronting epidemics of the modern era. *Int J Epidemiol* 2008; 37: 966–977.
82. Glanz K, Yaroch AL. Strategies for increasing fruit and vegetable intake in grocery stores and communities: policy, pricing, and environmental change. *Prev Med* 2004; 39(suppl 2): S75–S80.
83. Cervero R, Duncan M. Walking, bicycling, and urban landscapes: evidence from the San Francisco Bay Area. *Am J Public Health* 2003; 93: 1478–1483.
84. Pucher J, Dill J, Handy S. Infrastructure, programs, and policies to increase bicycling: an international review. *Prev Med* 2010; 50: S106–S125.
85. Tester J, Baker R. Making the playfields even: evaluating the impact of an environmental intervention on park use and physical activity. *Prev Med* 2009; 48: 316–320.
86. Smith KR, Brown BB, Yamada I, et al. Walkability and body mass index density, design, and new diversity measures. *Am J Prev Med* 2008; 35: 237–244.
87. Heindel JJ. Endocrine disruptors and the obesity epidemic. *Toxicol Sci* 2003; 76: 247–249.
88. Acerini CL, Hughes IA. Endocrine disrupting chemicals: a new and emerging public health problem? *Arch Dis Child* 2006; 91: 633–641.
89. Calafat AM, Ye X, Wong LY, et al. Exposure of the U.S. population to bisphenol A and 4-tertiary-octylphenol: 2003–2004. *Environ Health Perspect* 2008; 116: 39–44.
90. Sathyanarayana S, Karr CJ, Lozano P, et al. Baby care products: possible sources of infant phthalate exposure. *Pediatrics* 2008; 121: e260–e268.
91. Hatch EE, Nelson JW, Qureshi MM, et al. Association of urinary phthalate metabolite concentrations with body mass index and waist circumference: a cross-sectional study of NHANES data, 1999–2002. *Environ Health* 2008; 7: 27.
92. Trasande L, Landrigan PJ. The National Children's Study: a critical national investment. *Environ Health Perspect* 2004; 112: A789–A790.
93. Landrigan PJ, Trasande L, Thorpe LE, et al. The National Children's Study: a 21-year prospective study of 100,000 American children. *Pediatrics* 2006; 118: 2173–2186.
94. Landrigan PJ, Carlson JE, Bearer CF, et al. Children's health and the environment: a new agenda for prevention research. *Environ Health Perspect* 1998; 106(suppl 3): 787–794.
95. Food Quality Protection Act of 1996. http://www.epa.gov/opp00001/regulating/laws/fqpa/gpogate.pdf. Accessed January 2010.
96. Adler NE, Ostrove JM. Socioeconomic status and health: what we know and what we don't. *Ann N Y Acad Sci* 1999; 896: 3–15.

You Are Born with Your Genes: Justice and Protection against Discrimination in the Use of Genetic Information

Leslie P. Francis, PhD, JD

University of Utah, Salt Lake City, UT

OUTLINE

Medicine Today:
  Relevance (or Irrelevance) of Race to Genetic Information
Genetic Exceptionalism and Health Care
Protection From Genetic Discrimination: Where Scrutiny Matters
  Informed Consent to Genetic Testing
  Health Insurance Portability and Accountability Act
  Genetic Information Nondiscrimination Act
  Family Medical Leave Act
  Proposed Sequestration of Interoperable Electronic Records
Conclusion

ABSTRACT

Genetic information poses an apparent paradox for justice in health care. On the one hand, genetic information seems to be exactly the kind of information that it is unjust to take into account in the distribution of important goods. On the other hand, genetic information may increasingly become useful in individual patient care as evidence for personalized medicine increases. Although it is important to emphasize the difference between genetic information and the social construct of race, nonetheless, the historical interplay between discrimination based on assumptions about heritability and discrimination based on race are significant. Thus, this article defends the importance of paying very close attention to protections in the use of genetic information. Three current examples of inadequate protection are cited. First, informed consent to genetic testing should incorporate potential confidentiality risks raised by this information. Second, the Genetic Information Nondiscrimination Act is too porous to protect employees from their employers' possession of genetic information; if such information is used in discriminatory ways, burdens of seeking protection then fall on the employees. More stringent strategies for protecting information are thus desirable. Finally, current designs of interoperable health records fail to allow for the sequestration of genetic information. Efforts to develop such sequestration techniques should be encouraged; among other benefits, they would be helpful in shielding genetic information from employers. ***Mt Sinai J Med 77:188–196, 2010.*** *© 2010 Mount Sinai School of Medicine*

Address Correspondence to:

Leslie P. Francis
University of Utah
Salt Lake City, UT
Email: francisl@law.utah.edu

Key Words: discrimination, Genetic Information Nondiscrimination Act (GINA), genetics, justice.

Genetic information poses an apparent paradox for justice in health care. On the one hand, genetic information seems archetypical of the kind of information that it is unjust to take into account in distributing important goods such as health care. People do not control their genes (although their parents may make some choices that affect their genetic makeup), at least with today's technology; they are born with them. Genetic differences can be apparently random, and their consequences can range from minimal to devastating, sometimes unpredictably so. Long before any genetic explanations were understood, inherited differences were associated with differential treatment in myth, religion, and historical practice. Genetic determinism–the idea

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20170

© 2010 Mount Sinai School of Medicine

that genotype ineluctably results in phenotype—is empirically false and theoretically muddled. The interplay between racism and genetics in such practices as mandatory sterilization remains deeply troubling. Surely on the basis of these concerns, it would seem wrong for genetic differences to make any difference in the health care that people receive. On the other hand, in contemporary medicine, genetic (and now epigenetic) information seems increasingly clinically relevant to diagnosis and treatment. The burgeoning interest in personalized medicine, with genetic differences critical to assessing likely therapeutic responses, will only continue. Thus, clinical medicine seems exactly the arena in which the use of genetic information is most compelling.

> *Genetic information poses an apparent paradox for justice in health care. On the one hand, genetic information seems archetypical of the kind of information that it is unjust to take into account in distributing important goods such as health care. On the other hand, in contemporary medicine, genetic (and now epigenetic) information seems increasingly clinically relevant to diagnosis and treatment.*

What are we to make of this apparent paradox? Is clinical medicine just different, somehow immune from the mistakes and downright atrocities that have plagued the use of information about inheritance and genetics, and able to use such information for good but not for ill? Are the questions in the preceding paragraph merely addressing different kinds of issues, on the one hand policy issues such as access, costs, and confidentiality and on the other hand clinical management? If so, can the 2 kinds of issues be neatly separated, or will questions about clinical management ineluctably have implications for what care is funded or what information is kept confidential? This article argues that it is not unjust to take genetic information into account in clinical medicine, just as it is not unjust to take other physiological information into account when there is a sound scientific basis for doing so.

> *It is not unjust to take genetic information into account in clinical medicine, just as it is not unjust to take other physiological information into account when there is a sound scientific basis for doing so.*

Nor is genetic exceptionalism true in a straightforward way when health policy issues are raised; reasons that have been offered for thinking that genetic information is special also apply to other kinds of medical information. Nonetheless, the problems with genetic information add up for health policy in a way that problems with other types of physiological information may not, and they warrant social protections for genetic information even when it is used appropriately in health care treatment. A particularly important and troubling way in which they add up has implications for race: if race is sometimes used as a proxy for genetic information, the problems with genetic information may compound and entrench issues of racism.

Three protections for genetic information are especially important. First, patients should have the ability to consent specifically to genetic testing in a process that is as transparent as possible about any potential risks from uses of the information. Second, patients should be able to sequester certain categories of sensitive health information when interoperable electronic health record systems are networked for purposes of treatment, and genetic information should be among these designated categories. Third, although it is a step in the right direction, the Genetic Information Nondiscrimination Act (GINA) is insufficiently protective against both geneticism and racism.

MEDICINE TODAY: RELEVANCE (OR IRRELEVANCE) OF RACE TO GENETIC INFORMATION

The mapping of the human genome has provided a renewed impetus for the translation of genetic information into medical care.[1] Although the actual clinical results to date are sparse, hopes for pharmacogenetics remain high because of the assumption that drug metabolism and efficacy may vary genetically. One primary example is HER2/neu testing of metastatic breast cancer patients to determine responsiveness to Herceptin.[2] There are suggestions that drugs that were found to be ineffective or risky and thus were dropped from further study because of inefficacy or an adverse event profile might be resurrected for use in genetically defined subgroups. To date, however, payoffs for patient care remain limited.[3]

Within discussions of genetically personalized medicine, self-identified race has sometimes been used as a proxy for genotypic information. Thus used, however, race is a problematic proxy:

> Given the considerable genetic variation within conventional or self-identified racial and ethnic groups, attempts to use racial or ethnic designations as proxies for genetic variation are likely to be scientifically suboptimal and medically impractical. Such attempts can result in imprecise prescription guidelines and reinforce a public view of biologically defined races. As such, genotypic information should be taken into account along with other relevant patient information when prescribing medications.[2]

Within discussions of genetically personalized medicine, self-identified race has sometimes been used as a proxy for genotypic information. Thus used, however, race is a problematic proxy.

Parenthetically, although the term *personalized medicine* is used for genotypically tailored therapy, it is unclear why the term should be limited to this kind of individualized physiological feature of a person. Therapy tailored to an individual's history of toxic exposures would seem similarly personalized, for example.

The controversy over BiDil, a compound of hydralazine and isosorbide dinitrate cleared by the Food and Drug Administration for marketing for heart failure in patients who self-identify as black, is instructive.[4] Considerable scientific criticism has been directed against the retrospective race-based subgroup analyses performed on the original trials of BiDil. The subsequent African American Heart Failure Trial of BiDil has been criticized methodologically for relying on self-identification of race and for failing to compare responses in white patients.[5] A number of commentators have expressed the concern that the approval of this racialized therapy both signals and reinforces confusion between genetic difference and race as a socially constructed category. A serious risk is that it may direct attention away from the role of social factors in disease causation.[6,7]

Personalized medicine in the sense of genetically individualized therapy is practically and theoretically different from race-based therapy, however.[8] Nonetheless, given the use of self-identified race as a convenient proxy, if genetic information becomes increasingly important in medicine, it will be important to be vigilant to avoid the confusions of the BiDil experience. Obasogie[9] has therefore urged us to borrow the concept of strict scrutiny developed in constitutional law to examine any use of race identification in medical research, whether it is used as a surrogate for genetic difference or to address minority health concerns more generally. Such strict scrutiny demands that the case for using race must be compelling and narrowly tailored to the question at hand. If genetic information could be used instead of race as a proxy, for example, the test of strict scrutiny would not be met.

GENETIC EXCEPTIONALISM AND HEALTH CARE

Genetic exceptionalism is the idea that genetic information is a special type of health care information and hence should be given special treatment. A number of reasons have been offered for believing that genetic information is exceptional, and these range from the idea that genes are beyond individual control to the idea that genetic information is ineluctably familial.[10] Each of these arguments is overinclusive and thus not narrowly tailored to the use of genetic information, however: each applies to other kinds of medical information, too.[11] Although none of these reasons by itself singles out genetic information from other types of health care information, taken together, they provide a strong reason for particular care about the use of genetic information in medical care, even when the use is thought highly likely to be beneficial.

In standard liberal theories of justice, "people's opportunities should not be significantly limited due to factors that are wholly beyond their control."[12] If any features of persons are beyond their control, the characteristics that they receive from birth are, although of course to greater or lesser extents, people may be able to shape the effects that their genetic characteristics have on their lives. New developments in gene therapy may also eventually make genetic characteristics amenable to alteration. At present, however, genetic makeup is the paradigm case of a feature that is beyond individual control. Thus, on this basis, it would arguably be unjust to let genetic endowments determine opportunities, including access to health care; instead, a just society will work to make opportunities possible for people despite what may be difficulties resulting from their genetic inheritance.

Nor should genetic characteristics be assumed to determine phenotype. Naïve genetic determinism and the genetic essentialism that it fosters have long been the subjects of criticism.[13] That a trait is heritable means only that inheritance explains a significant percentage of its variance, not that it is genetically caused. For example, the presence of genes that affect skin color may explain a significant percentage of the variance in the frequency of hypertension without being directly causally linked to the hypertension. Instead, responses to people with the skin color trait may trigger biological responses in those people associated with stress, including effects on blood pressure.

In addition, the judgment that a trait is heritable bears no necessary link to the ease or difficulty in changing the trait. Changing the social responses that trigger stress, for example, may be extremely difficult in a society in which attitudes toward race remain intractable. By comparison, inheritance explains a great deal of the variety in myopia, but eyeglasses and laser-assisted in situ keratomileusis are relatively simple correctives in many people.[13] When social responses are particularly difficult to identify and much less change as responses associated with discrimination may be, the social protections discussed in this article are particularly important.

Tabery[14] has recently argued that the concept of a genetic predisposition also is muddled, confusing associations of a genetic characteristic with a phenotypic difference in a uniform direction (differences in scale) with associations that are not uniform (differences in rank). With respect to the latter, Tabery argues that the proper conceptualization is genetic interactionism, which captures the idea of susceptibility to a gene-environment interaction without the idea of "a gene for." An illustration of the importance of this distinction in behavioral genetics would be the discovery of an apparent association between the monoamine oxidase A (MAOA) gene, childhood abuse, and antisocial personality disorder. The variant of the gene that codes for low levels of MAOA was characterized erroneously as a gene for aggression, but now the actual association has been claimed to be a far more complex set of interactions: people with low levels who are not exposed to childhood abuse are less likely to be aggressive than people with higher levels who are not exposed to childhood abuse. In the presence of abuse, however, the relationship reverses. Such genetic interactionism, Tabery argues, complicates our ethical judgments because we cannot simply assume that an intervention will be best in all cases. Whether or not further research bears out the particulars of the MAOA example, the more general point remains: with gene-environment interactions, knowledge of genotype is not predictive of phenotype in the absence of information about the variety of gene-environment interactions and the likely presence of a given environment.

Another feature of genetic traits that has been used to argue that they warrant special treatment is that they are vertically transmitted from parent to child. However, so are some infectious diseases, such as congenital syphilis, herpes, or human immunodeficiency virus infection. Available preventive interventions may be different; for example, treatment of human immunodeficiency virus infection in a pregnant woman will reduce the likelihood of transmission to her offspring, whereas treatment of cystic fibrosis in a pregnant woman will not affect whether she transmits the genetic condition to her child. However, it does not follow from this difference that the types of vertical transmission warrant differential treatment for the factors involved.[11] Similarly, genetic information about one individual may give information about risks for family members. However, information about other factors, such as toxic exposures, may also provide such information if it is known that members of the family shared living conditions. As we come to know more about gene-environment interactions, moreover, differences between the familial significance of genetic information and the familial significance of toxic exposures may lessen still further.

Finally, and most importantly, there is a long history of stigmatization and discrimination associated with inherited traits and especially with race. Justice Holmes' infamous remark in an opinion for the US Supreme Court that "three generations of imbeciles are enough" has stood as a culminating symbol for this history.[15] Beliefs about racial differences in the heritability of intelligence have been used to justify power relationships[16] and even slavery and extermination.[17] From the misshapen profile in Edgar Degas' sketches of Miss La La[18] to the racist phrenology of 19th century anthropologists,[19] views about the biology and heritability of supposedly racial characteristics have left (mis)shapen but pervasive cultural attitudes about racial differences.

Each of these factors on its own might seem a limited or weak argument for regarding genetic information as special. For example, Posner[20] makes this argument in criticism of singling out genetic information for protection against discrimination in health insurance:

> There is no morally relevant distinction between looking at the person's outer shell and looking at his blood under a microscope.

Posner is correct that ethical lines do not cut neatly between genetic information and other types of biological information, as outlined previously. If we lived in a world in which people had a proper understanding of the significance of genetic information and in which the history of discrimination based on inherited characteristics intertwined with discrimination based on race, we could simply look at biological characteristics through clear lenses. However, we live in a world of unclear vision in which the mistakes that have been associated with genetic information are of great moral importance. Posner's argument leaves aside the mistaken assumptions about genetic determinism, about the nature of gene-environment interactions, and about heritability and race, which have deeply corrupted the social impact of this information.

Moreover, when a confluence of social factors is in play that risks making the most vulnerable even worse off, as they are here, there is special reason for concern from the point of view of social justice. Attention to the situation of those who are already in circumstances of injustice is a core problem for partial compliance theory, that is, for theorizing about justice under circumstances of injustice.[21] Arguably, the interplay between genetic information and race poses exactly this kind of partial compliance problem as mistaken beliefs about inheritance and race may deepen discrimination.[13] Fears on the part of the public about employers or insurers acquiring their genetic information–fears that are greater than those about other types of health information–thus may be reasonable under current circumstances, even though they might not be in a world less inflected by histories of discrimination.[22,23] This concern is also the reply to commentators such as Rothstein[24] who are concerned that singling out genetic information for special treatment may become a self-fulfilling prophecy that increases the risks of stigmatization. Thus, as Obasogie[9] argues, the greatest care must be taken before genetic information is put into play in medical care or in health policy.

PROTECTION FROM GENETIC DISCRIMINATION: WHERE SCRUTINY MATTERS

This section briefly explores 3 areas in which special care may be warranted with respect to the use of genetic information in health care: informed consent to genetic testing; the limits of GINA protection; and the possibility of designing interoperable electronic health records allowing the designation of certain categories of sensitive health information, including genetic information.

Informed Consent to Genetic Testing

First, it must be recognized that informed consent to genetic testing is a complex matter. Where testing may indicate disease risk, consent should include not only the medical risks and benefits of the testing (including what is or is not known about the accuracy of the test) but also information about how test results will be recorded and protected.

Informed consent to genetic testing is a complex matter. Where testing may indicate disease risk, consent should include not only the medical risks and benefits of the testing (including what is or is not known about the accuracy of the test) but also information about how test results will be recorded and protected.

As different types of genetic testing and uses of genetic information come into wider medical use (eg, through personalized medicine), informed consent will need to explore the potential risks and benefits of this information for patients. For example, the inclusion of genetic information in an interoperable electronic health record means that the information, in the absence of any special protection, will go wherever that electronic health record is transmitted. Patients should be aware then that if their electronic records are accessible by other providers for purposes of treatment, as they surely will be, the genetic information will be included.

Health Insurance Portability and Accountability Act

Patients should also be aware that if they authorize release of their health information pursuant to the Health Insurance Portability and Accountability Act (HIPAA) of 1996, their genetic test information will

Patients should also be aware that if they authorize release of their health information pursuant to the Health Insurance Portability and Accountability Act (HIPAA) of 1996, their genetic test information will be included in the absence of any special protections.

be included in the absence of any special protections. A particular question that is as yet not adequately addressed by HIPAA is the relevance of genetic information to the creation of de-identified data sets that may be used in research. Once data have been de-identified, HIPAA protections, including the requirement for patient authorization for data release, no longer apply. Data are de-identified under the HIPAA safe harbor rule when each of a list of 19 identifying factors has been removed; this list includes biometric identifying information (the examples given are finger and voice prints) and any other information that there is reason to believe will be uniquely identifying.[25] Although inclusion of a fully sequenced genome would clearly be uniquely identifying, questions about how much genetic information raises risks of re-identification remain incompletely examined. For example, it has recently been argued that individuals can be re-identified in certain genome-wide association studies.[26] Given these uncertainties, it is especially important that patients be informed in advance if there is any possibility that their genetic test results might be included in informational research studies, even as part of a de-identified data set, and that they be asked whether they would be willing to permit use of their data without explicit consent to each study. Similar uncertainties attend the protection of stored tissue samples in biobanks.[27]

Beyond questions about the efficacy of de-identification efforts, patients may also have cultural or other reasons for wishing to be given the opportunity to consent to any use of their genetic information in research. Concerns about group stigmatization are also important. The cultural objections of members of the Havasupai to uses of their genetic information in research involving migration patterns are by now well known.[28]

Genetic Information Nondiscrimination Act

Second, GINA prohibits health insurers or employers from making certain decisions on the basis of genetic information. Health insurers offering either group or individual plans may not use genetic information to determine eligibility or to underwrite premiums.[29] Proposed rules published in October 2009 would amend the HIPAA privacy rule to restrict the ability of covered entities to use genetic information to set group premium levels.[30] Also under GINA, employers may not seek or use genetic information in making employment decisions.[31] Finally, GINA protection extends to uses of genetic information, but GINA does not protect people from what may be regarded as discriminatory uses of actual health conditions.[32]

With respect to genetic information, GINA is not a panacea, however. When we consider the statute's likely effect, it is important to understand the difference between a confidentiality statute (a statute such as HIPAA that restricts information flow) and a nondiscrimination statute (a statute with the goal of protecting people against discrimination). There is a sense in which confidentiality statutes are more protective than nondiscrimination statutes, as the former may ensure that information that is a potential source of discrimination never falls into the hands of those who are likely to discriminate. Nondiscrimination statutes, in contrast, protect people by prohibiting discrimination and creating legal causes of action challenging allegedly discriminatory actions that have already caused harm. The practical result of nondiscrimination statutes then is that they place the burden on people who have experienced discrimination to challenge it after the fact.[33,34]

GINA prohibits health insurers or employers from making certain decisions on the basis of genetic information but GINA is not a panacea placing the burden on people who have experienced discrimination to challenge it after the fact.

In GINA, there are important exceptions to the prohibitions on collecting genetic information that argue that the statute functions primarily as a nondiscrimination statute, despite the recognition on the part of Congress that protecting confidentiality would provide a more secure footing against discrimination.[35] Employers may gain access to genetic information when the request for the information is inadvertent.[36] This has been dubbed the water cooler exception to GINA, by which information about genetic testing or family medical history is acquired by the employer through casual conversation. Much genetic information may come into the hands of employers in this way if employees are not careful; this possibility should be widely publicized, and health care providers should be aware of the importance of reminding patients of this in the process of informed consent to genetic testing.

Family Medical Leave Act

A second, very significant exception to the prohibition on employers' collecting genetic information is the Family Medical Leave Act (FMLA). When an employee requests leave under FMLA for his or her own medical treatment or to care for a family member, GINA permits the employer to request the employee's medical history for verification of the need for the leave.[37]

> *When an employee requests leave under FMLA for his or her own medical treatment or to care for a family member, GINA permits the employer to request the employee's medical history for verification of the need for the leave.*

If genetic information is included in the request, employers will have it, and the result may be that they then use it in discriminatory fashion, although they will remain obligated to treat the information with the confidentiality requirements attendant on any other medical information in their possession. One example would be an employee who requests medical leave for treatment of her own breast cancer; a full medical record might include the results of any applicable genetic testing, yet those results would not be necessary to verify the need for the treatment of the manifested health condition. In this case, GINA would permit the employer to consider the impact of the employee's manifested health condition, breast cancer, as job-related, but the employer would not be permitted to consider any beliefs about the relevance (or irrelevance) of the genetic information to the possibility that the employee might develop other conditions. Another example would be a situation in which an employee requests FMLA leave to care for a relative with breast cancer and the employer gains knowledge of genetic information from the substantiating medical record; the employer in this case might make assumptions about the employee in the absence of any indication whatsoever of manifested disease. Similar problems might occur if an employee requests leave to care for a dependent child and the employer learns not only that the child has a physical disease requiring care but also that the disease is thought to have, at least in part, a genetic explanation. The likelihood that medical histories might be requested for FMLA verification is one of many reasons to support the idea that health records should be designed to permit patients to mask sensitive categories of health information before they are transferred. Rothstein[38] has made this same point in a parallel context: employers are permitted under the Americans with Disabilities Ace to request health records to determine fitness for duty once an offer for a position has been made.

Proposed Sequestration of Interoperable Electronic Records

Third, a response to concerns about the transmission and use by employers or others when genetic information is included in interoperable electronic health records is to provide for the sequestration of certain defined categories of sensitive information, including genetic information. The National Committee on Vital and Health Statistics[39] has proposed that, as interoperable electronic records on a wide scale are developed, they should include the functionality to mask such predetermined categories of sensitive information.

> *The National Committee on Vital and Health Statistics has proposed that, as interoperable electronic records on a wide scale are developed, they should include the functionality to mask such predetermined categories of sensitive information.*

The National Committee on Vital and Health Statistics proposal provides for a break-the-glass feature in cases of emergency treatment when the patient cannot give consent and for flagging notation if it is necessary to indicate that sequestration has occurred. This design feature would allow patients to request the masking of genetic information when transfers occur for research or in response to FMLA requests or in the case of other patient-authorized transfers. Without such a design feature, such requests would be difficult to implement. Interoperable record design is at present in the development phase, and it is thus particularly important at this juncture to explore design capacities for such sequestration.

CONCLUSION

The misunderstanding and misuse of supposed information about inheritance historically and of genetic information more recently pose an apparent paradox for contemporary health care. In some circumstances, there may be an appropriate evidence basis for the use of such information in clinical care. Yet genetic information, or the use of proxies for it, seems exactly the kind of information that it is inappropriate to use as a basis for differential treatment more generally. In order to dissolve this paradox, it is important to subject any use of genetic information to particularly careful scrutiny and management. Three particular management techniques have been defended here: the nature of informed consent to genetic testing, the recognition that GINA provides at best imperfect protection against genetic discrimination, and the effort to design interoperable electronic health records to permit sequestration of a defined category of genetic information upon a patient's request.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Secretary's Advisory Committee on Genetics, Health and Society. The integration of genetic technologies into health care and public health. http://oba.od.nih.gov/oba/SACGHS/SACGHS%20Progress%20and%20Priorities%20Report%20to%20HHS%20Secretary%20Jan%202009.pdf. Published January 2009. Accessed January 2010.
2. Secretary's Advisory Committee on Genetics, Health and Society. Realizing the potential of pharmacogenomics: opportunities and challenges. http://oba.od.nih.gov/oba/SACGHS/reports/SACGHS_PGx_report.pdf. Published May 2008. Accessed January 2010.
3. Genetics and Public Policy Center. Issue brief: pharmacogenomics: hope and hype. http://www.dnapolicy.org/policy.issue.php?action=detail&issuebrief_id=47. Published September 2008. Accessed January 2010.
4. Kahn J. BiDil and racialized medicine. http://www.councilforresponsiblegenetics.org/pageDocuments/PLMVM6FTAO.pdf. Accessed January 2010.
5. Ellison GTH, Kaufman JS, Head RF, et al. Flaws in the U.S. Food and Drug Administration's rationale for supporting the development and approval of BiDil as a treatment for heart failure only in black patients. *J Law Med Ethics* 2008; 36: 449–457.
6. Roberts D. Is race-based medicine good for us? African American approaches to race, biomedicine, and equality. *J Law Med Ethics* 2008; 36: 537–545.
7. Duster T. Race and reification in science. *Science* 2005; 307: 1050–1051.
8. Tutton R, Smart A, Martin PA, et al. Genotyping the future: scientists' expectations about race/ethnicity after BiDil. *J Law Med Ethics* 2008 36: 464–470.
9. Obasogie OK. Beyond best practices: strict scrutiny as a regulatory model for race-specific medicines. *J Law Med Ethics* 2008; 36: 491–497.
10. Secretary's Advisory Committee on Genetic Testing. Enhancing the oversight of genetic tests: recommendations of the SACGT. http://oba.od.nih.gov/oba/sacgt/reports/oversight_report.pdf. Published July 2000. Accessed January 2010.
11. Battin MP, Francis LP, Jacobson JA, Smith CB. *The Patient as Victim and Vector. Ethics and Infectious Disease*. New York, NY: Oxford University Press; 2009; 184–199.
12. Buchanan A, Brock DW, Daniels N, Wikler D. *From Chance to Choice: Genetics and Justice*. Cambridge, United Kingdom: Cambridge University Press; 2000; 17.
13. Parens E. Genetic differences and human identities: why talking about behavioral genetics is important and difficult. *Hastings Cent Report* 2004; 34: S1–S36.
14. Tabery J. From a genetic predisposition to an interactive predisposition: rethinking the ethical implications of screening for gene-environment interactions. *J Med Philos* 2009; 34: 27–48.
15. *Buck v Bell*, 274 US 200, 207 (1927).
16. Shklar J. Redeeming American political theory. *Am Pol Sci Rev* 1991; 85: 3–15.
17. Ossario P. Race, genes and intelligence. http://www.councilforresponsiblegenetics.org/pageDocuments/XT0UW7SVHY.pdf. Accessed January 2010.
18. Brown MR. Miss La La's teeth: reflections on Degas and race. *Art Bull* 2007; 89: 738–765.
19. Harrison FV. The persistent power of "race" in the cultural and political economy of racism. *Annu Rev Anthropol* 1995; 24: 47–71.
20. Posner E. The puzzling consensus in favor of the Genetic Information Nondiscrimination Act 2009. http://www.slate.com/blogs/blogs/convictions/archive/2008/05/04/the-puzzling-consensus-in-favor-of-the-genetic-information-nondiscrimination-act.aspx. Accessed January 2010.
21. Rawls J. *A Theory of Justice*. Boston, MA: Belknap Press; 1971.
22. Schlein D. New frontiers for genetic privacy law: the Genetic Information Nondiscrimination Act of 2008. *George Mason Univ Civil Rights Law J* 2009; 19: 311–369.
23. Silvers A, Stein MA. Human rights and genetic discrimination: protecting genomics' promise for public health. *J Law Med Ethics* 2003; 31: 377–389.

24. Rothstein M. GINA's beauty is only skin deep. http://www.councilforresponsiblegenetics.org/GeneWatch/GeneWatchPage.aspx?pageId=184. Accessed January 2010.
25. 45 CFR §164.514(b)(2)(i)(P),(R) (2009).
26. Homer N, Szelinger S, Redman M, et al. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet* 2009; 4: e1000167.
27. Andrews LB. Harnessing the benefits of biobanks. *J Law Med Ethics* 2005; 33: 22–28.
28. McGregor J. Population genomics and research with socially identifiable groups. *J Law Med Ethics* 2007; 35: 356–370.
29. Publ L No. 110–223, 122 Stat 881, Title I (2008).
30. *Fed Regist* 2009; 74: 51698–51710.
31. Publ L 110–223, 122 Stat 881, Title II (2008).
32. McGuire AL, Majumder MA. Two cheers for GINA? *Genome Med* 2009; 1: 6.
33. Colker R. Winning and losing under the Americans with Disabilities Act. *Ohio State Law J* 2001; 62: 239–279.
34. Colker R. The Americans with Disabilities Act: a windfall for defendants. *Harv Civil Rights Civil Liberties Law Rev* 1999; 34: 99–162.
35. Snowe O. Introductory remarks of Senator Olympia Snowe (R-ME) on the Genetic Information Nondiscrimination Act–January 22, 2007. http://www.geneticalliance.org/ksc_assets/publicpolicy/gina_remarks_snowe.pdf. Accessed January 2010.
36. Equal Employment Opportunity Commission. Regulations under the Genetic Information Nondiscrimination Act of 2008. *Fed Regist* 2009; 74: 9061.
37. 42 USC §2000 ff-1(b)(3) (2009).
38. Rothstein M. Occupational health and discrimination issues raised by toxicogenomics in the workplace. In: Sharp RR, Marchant GE, Grodsky JA, eds. *Genomics and Environmental Regulation: Science, Ethics, and Law*. Baltimore, MD: Johns Hopkins University Press; 2008; 183–201.
39. National Committee on Vital and Health Statistics. Individual control of sensitive health information available over the National Health Information Network for purposes of treatment. http://www.ncvhs.hhs.gov/080220lt.pdf. Published February 2008. Accessed January 2010.

SPECIAL FEATURE

Type 2 Diabetes and Cancer: What Is the Connection?

Dara Cannata, Yvonne Fierz, Archana Vijayakumar, and Derek LeRoith

Mount Sinai School of Medicine New York, NY

OUTLINE

INSULIN AND INSULIN-LIKE GROWTH FACTORS
OVEREXPRESSION OF INSULIN RECEPTOR AND INSULIN-LIKE GROWTH FACTOR I RECEPTOR IN CANCER
DIABETES AND SITE-SPECIFIC CANCERS
- **Breast Cancer**
- **Colorectal Cancer**
- **Pancreatic Cancer**
- **Prostate Cancer**

EFFECTS OF MEDICATIONS ON THE ASSOCIATION BETWEEN TYPE 2 DIABETES MELLITUS AND CANCER
- **Sulfonylureas**
- **Insulin**
- **Metformin**
- **Thiazolidinediones**

CONCLUSION

ABSTRACT

Epidemiological studies have demonstrated an association between type 2 diabetes and cancer. Type 2 diabetes is characterized by insulin resistance and hyperinsulinemia. Hyperinsulinemia may lead to cancer through insulin's effect on its cognate receptor and the insulin-like growth factor system. The effects of insulin and insulin-like growth factor I on cancer development and progression have been demonstrated in animal and human studies. Type 2 diabetes has been positively associated with cancers of the breast, colon, and pancreas. An inverse relationship has been observed between type 2 diabetes and prostate cancer, and this may be due to lower testosterone levels in men with type 2 diabetes. Medications used to treat type 2 diabetes may affect cancer cells directly or indirectly by affecting serum insulin levels. Hyperinsulinemia may be an important risk factor for cancer as well as a target for cancer therapy. ***Mt Sinai J Med 77:197–213, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: cancer, diabetes, hyperinsulinemia, insulin resistance, insulin-like growth factor.

Address Correspondence to:

Derek LeRoith
Mount Sinai School of Medicine
New York, NY
Email: derek.leroith@mssm.edu

Diabetes mellitus is composed of a group of disease entities marked by hyperglycemia resulting from abnormalities in insulin production and/or action. In 2008, the United States alone was estimated to have 23.6 million people with diabetes.[1] Approximately 1.6 million new cases of diabetes were diagnosed in people 20 years of age or older in 2007, and this figure has steadily risen over the past few decades.[1] The increase in the incidence of diabetes in adults is thought to be due mainly to type 2 diabetes (T2D), which accounts for 90% to 95% of all diagnosed cases of diabetes and is typified by insulin resistance, hyperinsulinemia, and hyperglycemia. In the insulin-resistant state, there is reduced responsiveness of the skeletal muscle, liver, and adipose tissue to insulin. As a result, insulin levels rise in an attempt to maintain euglycemia.[2] As insulin resistance becomes severe, overt hyperglycemia develops. Epidemiological studies have demonstrated increased risk of and mortality from many different types of cancer, including cancers of the colon, pancreas, and female breast (Tables 1 and 2). As noted in table, relative risk for five case studies is 1.18.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20167

© 2010 Mount Sinai School of Medicine

Table 1. *Diabetes and Relative Risk of Cancer By Site.*

Cancer	Meta-Analysis	Relative Risk (95% Confidence Interval)
Breast	Larsson *et al.*[43] (2007)	
	5 case control studies	1.18 (1.05–1.32)
	15 cohort studies	1.20 (1.11–1.30)
Colon	Larsson *et al.*[65] (2005)	
	6 case control studies	1.36 (1.23–1.50)
	9 cohort studies	1.29 (1.16–1.43)
Pancreas	Huxley *et al.*[80] (2005)	
	17 case control studies	1.94 (1.53–2.46)
	19 cohort studies	1.73 (1.59–1.88)
Prostate	Kasper and Giovannucci[91] (2006)	
	9 case control studies	0.89 (0.72–1.11)
	10 cohort studies	0.81 (0.71–0.92)

Table 2. *Diabetes and Relative Risk of Cancer Mortality by Site.*

Cancer	Studies	Relative Risk (95% Confidence Interval)
Breast	Coughlin *et al.*[54] (2004)	1.27 (1.11–1.45)
	Verlato *et al.*[53] (2003)	1.40 (1.06–1.81)
	Yancik *et al.*[52] (2001)	1.76 (1.23–2.52)
Colon	Larsson *et al.*[65] (2005)	1.26 (1.05–1.50)
	Coughlin *et al.*[54] (2004)	Male: 1.20 (1.06–1.37); female: 1.24 (1.07–1.43)
Pancreas	Coughlin *et al.*[54] (2004)	Male: 1.48 (1.27–1.73); female: 1.44 (1.21–1.72)
Prostate	Coughlin *et al.*[54] (2004)	0.90 (0.80–1.02)
	Ma *et al.*[98] (2008)*	2.38 (1.31–4.30)†

*Ma *et al.* studied the relationship between the plasma C-peptide concentration and prostate cancer–specific mortality.

† Hazard ratio.

Epidemiological studies have demonstrated that patients with diabetes mellitus have increased risk of and mortality from many different types of cancer, including cancers of the colon, pancreas, and female breast.

T2D is commonly comorbid with obesity, which has also been linked to cancer.[3] However, higher rates of cancer and cancer-related mortality have been seen in T2D patients, regardless of the body mass index, and this suggests that T2D may be an independent risk factor for cancer and cancer-related mortality.[4,5] The major factors implicated in the tumor-promoting effect of T2D are chronic hyperinsulinemia and hyperglycemia.[6]

Many studies exploring the relationship between diabetes and cancer do not distinguish between type 1 diabetes (T1D) and T2D. T1D is an autoimmune disease that arises because of pancreatic β cell destruction with resultant insulin deficiency and hyperglycemia. On the other hand, T2D develops because of peripheral insulin resistance and is usually characterized by β cell hyperplasia and hyperinsulinemia. The 2 diseases also differ in their age of onset; although T1D is usually diagnosed in juvenile patients, T2D occurs mostly in adults. Given the differences in the pathophysiology of the 2 conditions, it is important to make this diagnostic distinction before any conclusions are made about the association between diabetes and cancer. However, given that the large majority of diabetic patients have T2D and that most studies have been conducted on patients diagnosed with diabetes at an older age, one could extrapolate that the majority of diabetic patients who develop cancer are T2D patients. Data from studies of the association between T1D and cancer are sparse in comparison with data on T2D, and they contain fewer patients because of the smaller cohort of patients with T1D. Zendehdel and colleagues[7] observed cancer rates in almost 30,000 Swedish T1D patients and found no increased risk of breast, pancreatic, or colorectal cancer. Altogether, it is important to distinguish between T1D and T2D as they have different pathological mechanisms that may affect their associations with cancer. T2D, in which insulin resistance and hyperinsulinemia are present, has been consistently associated with cancer, whereas T1D, a disease of insulin deficiency, has not.

It is important to distinguish between T1D and T2D as they have different pathological mechanisms that may affect their associations with cancer. T2D, in which insulin resistance and hyperinsulinemia are present, has been consistently associated with cancer, whereas T1D, a disease of insulin deficiency, has not.

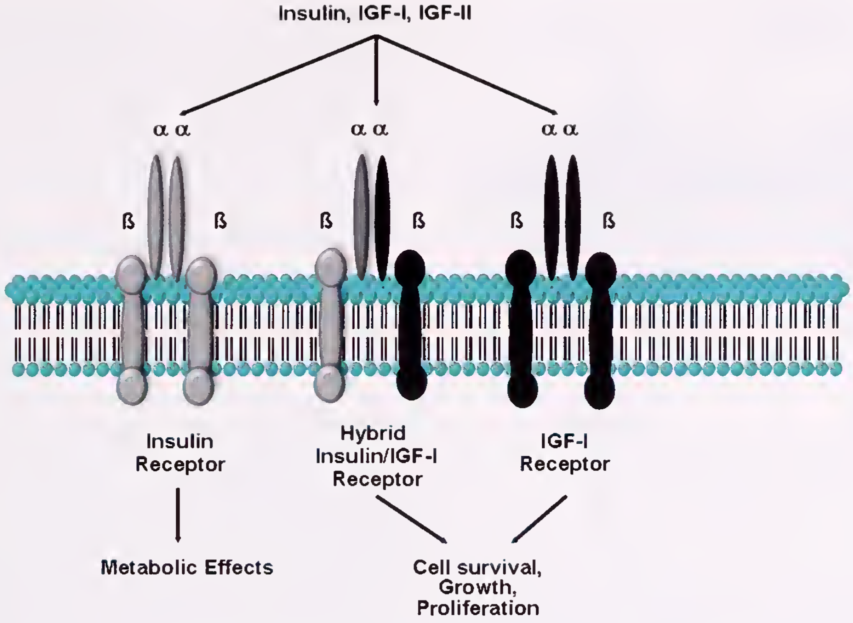


Fig 1. IR, IGF-IR, and hybrid receptors. IR and IGF-IR are transmembrane tyrosine kinase receptors. They are heterotetrameric proteins that contain 2 extracellular α subunits and 2 transmembrane β subunits. IR and IGF-IR can form hybrid receptors consisting of 1 α subunit and 1 β subunit of IR and 1 α subunit and 1 β subunit of IGF-IR. **Abbreviations:** IGF, insulin-like growth factor; IGF-IR, insulin-like growth factor I receptor; IR, insulin receptor.

Although the increase in the risk of cancer and cancer-related mortality attributed to T2D is relatively small, the fact that nearly 23.6 million Americans have T2D makes this a significant public health issue. This review further explores the association between T2D and cancer.

INSULIN AND INSULIN-LIKE GROWTH FACTORS

In order to examine the relationship between T2D and cancer, it is important to describe the insulin/insulin-like growth factor I (IGF-I) system and the role that it plays in the development and progression of cancer. Insulin and IGF-I are growth factors that are involved in glucose metabolism and in the activation of cellular signaling cascades mediating cell proliferation, differentiation, and survival.[8] In this section, we explore the insulin/IGF system and its link to cancer. Traditionally, insulin has been thought to have a metabolic effect, whereas IGF-I has been thought to have a mitogenic effect. However, as we discuss in this section, studies suggest that insulin may play a mitogenic role as well.

The IGF system is highly intricate and includes the ligands IGF-I, IGF-II, and insulin; their receptors insulin-like growth factor I receptor (IGF-IR), insulin-like growth factor II receptor (IGF-IIR), and insulin receptor (IR); insulin-like growth factor binding proteins (IGFBPs); and effector molecules downstream of IGF-IR and IR. The downstream cascade triggered by IGF-I and insulin involves effector molecules and pathways such as the IR substrate proteins, the phosphatidylinositol 3-kinase (PI3K) pathway, and the mitogen-activated protein kinase pathway.[9] PI3K activates the protein kinase B pathway.[10] The PI3K/protein kinase B pathway and the mitogen-activated protein kinase pathways promote cell growth and have antiapoptotic effects.[11,12] Both of these pathways play important roles in tumorigenesis.[13,14] IGF-IR and IR are structurally homologous. They are both heterotetrameric proteins (with 2 extracellular α subunits and 2 transmembrane β subunits) that possess intrinsic tyrosine kinase activity (Figure 1). Despite their shared homology and common signaling pathways, IR and IGF-IR have evolved to perform different functions and do so by activating and recruiting different docking proteins and signaling pathways.[11,15] IR is primarily involved in glucose homeostasis, whereas IGF-IR is involved in body growth through the effects of IGF-I on cell proliferation, differentiation, and apoptosis. IR is preferentially expressed in adipose, liver, and muscle tissues, whereas IGF-IR is expressed in almost all body tissues.[11]

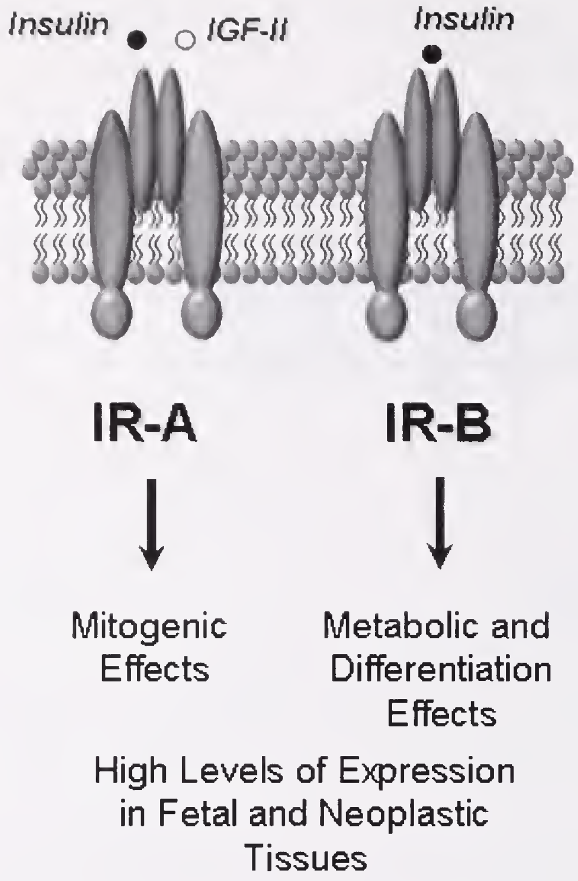

Insulin
IGF-II
Insulin
IR-A
IR-B
Mitogenic Effects
Metabolic and Differentiation Effects
High Levels of Expression in Fetal and Neoplastic Tissues

Fig 2. IR isoforms. IR is expressed as 2 isoforms, IR-A and IR-B, which are formed by alternative splicing of exon 11. IR-A is expressed in high levels in fetal and neoplastic tissues, binds insulin and IGF-II, and mediates mitogenic effects. IR-B predominantly binds insulin and exerts metabolic effects. IR-B is expressed in tissues on which insulin acts to exert its metabolic effects, such as the liver, muscles, adipocytes, and kidneys. In addition, IR-B has differentiation effects. **Abbreviations:** IGF, insulin-like growth factor; IR, insulin receptor.

IR has 2 isoforms, IR-A and IR-B, which are formed by alternative splicing of the IR transcript; IR-B contains exon 11, whereas IR-A does not.[16] In vivo studies have shown that the 2 isoforms have different effects on cells (Figure 2).[17] IR-A shows high levels of expression in fetal tissue.[18] IR-A mediates antiapoptotic and mitogenic effects, whereas IR-B has metabolic effects.[17] IR-A binds IGF-II and insulin, whereas IR-B is poorly responsive to IGF-II.[17] IGF-II is a powerful stimulator of cell proliferation.[18] Thus, the ability of IR-A to bind IGF-II may explain the difference in cell signaling by the 2 isoforms.[17–19]

Insulin and IGF-I bind their own cognate receptors with highest affinity.[20] Because of their high degree of homology, IGF-IR and IR can form hybrid receptors consisting of 1 α subunit and 1 β subunit of IR and 1 α subunit and 1 β subunit of IGF-IR (Figure 1). Hybrid receptors behave more like IGF-IR than IR as they have a higher affinity for IGF-I than insulin.[21,22] As a result, hybrid receptors are thought to provide additional binding sites for IGF-I and lead to increased cell sensitivity to IGF-I.[21,23]

OVEREXPRESSION OF INSULIN RECEPTOR AND INSULIN-LIKE GROWTH FACTOR I RECEPTOR IN CANCER

Both IGF-IR and IR have been found to be overexpressed in cancer cells.[18,24–28] Although IGF-IR overexpression is seen in many cancer cells, studies have yielded mixed results on the clinical and prognostic significance of this finding.[25,29,30] Papa *et al.*[25] reported that the average IGF-IR level in breast cancer tissue specimens was nearly 10-fold higher than that seen in normal breast tissue; however, elevated IGF-IR was associated with a favorable prognosis. On the other hand, multiple groups of researchers have found elevated expression levels of IGF-IR to be associated with worse prognosis in cancer.[29,30] In vivo studies have demonstrated that tumor suppressor genes, such as p53, p63, p73, breast cancer 1 early onset, von Hippel-Lindau protein, Wilms' tumor protein 1, and Kruppel-like factor 6, inhibit IGF-IR gene expression.[31–33] Thus, mutations of these tumor suppressor genes may lead to tumor growth and proliferation via a lack of inhibition of IGF-IR. In addition, mutations of these tumor suppressor genes lead to enhanced expression of IGF-IR.[31–33]

Increased IR levels have been seen in cancers of the breast, lung, ovary, thyroid, and colon.[18,24,26,27] Suggested mechanisms for IR overexpression in cancer include but are not limited to p53 inactivation (as p53 may down-regulate IR) and Wnt-1, Neu, and Ret oncogene activity.[34,35] Papa *et al.*[24] found that 80% of human breast cancers (ductal, papillary, mixed, lobular, and other subtypes) have elevated IR expression, some with IR levels 10-fold higher than the levels observed in normal breast tissue. Additionally, a high IR level has been associated with larger, higher grade tumors and higher estrogen receptor concentrations.[24] Mathieu and colleagues[36] studied tumor specimens from patients with node-negative breast carcinoma and found that patients with tumors with undetectable IR levels in malignant epithelial cells had lower predicted 5-year disease-free survival than patients with tumors with detectable IR. However, in a small subset of patients with very high IR levels, disease-free survival was decreased in comparison with that of patients with moderate IR expression. Thus, much like that of IGF-IR overexpression, the clinical significance of IR overexpression is not yet known. A recent study by Law *et al.*[37] may shed some light. This group showed that activated (phosphorylated) IGF-IR/IR could be detected in almost half of the breast tumors of patients with invasive breast cancer (diabetic status not specified). In addition, the total levels of IR and phosphorylated IGF-IR/IR were indicative of poor survival in these patients. As the authors did not differentiate between the activated IR, IGF-IR, and hybrid receptors, the relative importance of the activation of each receptor could not be stated. However, this study shows that the insulin/IGF-IR system plays a role in breast cancer growth and promotion. In addition, phosphorylated IGF-IR/IR may be more indicative of a poor outcome than the total IGF-IR/IR level, and this may explain why studies thus far have yielded conflicting results.

As stated previously, IR has 2 known isoforms; IR-A has predominantly mitogenic effects, whereas IR-B has effects that are mainly metabolic.[17] In vitro studies have shown that cancer cells possess an increased ability to express IR-A.[19,36,38] It has been demonstrated that IR-A is overexpressed in breast and colon tumors.[18,19] As cancer cells also demonstrate IR overexpression, they have high levels of IR-A.[18] In vitro studies have confirmed that elevated IR-A levels are associated with tumor progression.[27]

The overexpression of IR and IGF-IR in cancer cells can lead to increased expression of the hybrid receptors.[11] Many cancer tissues have increased levels of hybrid receptors.[21,39] For example, in both breast and thyroid carcinomas, studies have demonstrated a higher level of hybrid receptors in comparison with IGF-IR.[21,39] Therefore, hybrid receptors appear to play a crucial role in mediating the effects of IGF-I on cancer cells.

In T2D, chronic hyperinsulinemia may affect cancer risk not only directly through insulin's mitogenic potential but also indirectly through the more potent effects of IGF-I.

Aside from forming hybrid receptors, the insulin and IGF systems interact in several ways. Insulin hinders the production of 2 IGFBPs, IGFBP-1 and IGFBP-2, which transport IGF-I to the peripheral tissues.[4] Chronically elevated insulin levels, therefore, may increase bioavailable IGF-I.[4] In addition, when insulin levels are elevated, insulin is more likely to bind to IGF-IR, which has increased mitogenic potential in comparison with IR.[40]

Altogether, the association between T2D and cancer is mediated in part by chronic hyperinsulinemia acting through insulin's cognate receptor as well as the effects that elevated insulin levels have on the IGF system.

The association between T2D and cancer is mediated in part by chronic hyperinsulinemia acting through insulin's cognate receptor as well as the effects that elevated insulin levels have on the IGF system.

DIABETES AND SITE-SPECIFIC CANCERS

T2D has been linked to many different types of cancer. Here we discuss some of the most common cancer sites and their association with T2D.

Breast Cancer

Breast cancer was estimated to be responsible for 26% of cancer cases and 15% of cancer deaths in women in 2008.[41] Risks for breast cancer include age greater than 50 years, female gender, Caucasian and African American race, family history, and genetics.[42] Hormonal factors such as the duration of menstrual history, use of hormone replacement therapy, and nulliparity are also thought to play a role.[42] The link between T2D and breast cancer has been extensively studied, and the two are thought to be linked via activation of the insulin/IGF pathway through hyperinsulinemia and via dysregulation of sex hormones.[5,43] Cell proliferation of normal human breast tissue and breast cancer cell lines has been shown to be enhanced by insulin.[44,45] Insulin stimulated cell cycle progression and DNA synthesis of MCF-7 breast cancer cells in vitro.[45] Animal studies also have demonstrated insulin's mitogenic effects on breast tumor growth.[46,47] A-ZIP/F-1 transgenic mice have been used as a model of lipoatrophic T2D. A-ZIP mice have no white fat, reduced brown fat, and elevated glucose, insulin, free fatty acids, and triglycerides and also exhibit increased levels of inflammatory cytokines. The crossing of A-ZIP/F-1 mice with the C3(1)/T-Ag mammary tumor model demonstrated increased tumor incidence and multiplicity and decreased tumor latency in comparison with control mice.[48] We have developed the MKR model of T2D and studied the effect of hyperinsulinemia on breast cancer. In this transgenic mouse model, dominant negative IGF-I receptors are expressed specifically in skeletal muscle under the control of the muscle creatine kinase (MCK) promoter. The MKR mouse is insulin-resistant, glucose-intolerant, and hyperinsulinemic but not obese. Female MKR mice orthotopically inoculated with mouse mammary carcinoma cells have accelerated mammary tumor growth in comparison with nondiabetic control mice. This effect is abrogated when these mice are treated with a small-molecule IR/IGF-IR tyrosine kinase inhibitor, and this suggests that hyperinsulinemia is mediating the accelerated tumor growth in MKR mice via activation of IR and/or IGF-IR.[49]

Similar to the aforementioned mouse models, human studies have also demonstrated an association between insulin levels and breast cancer risk. Gunter *et al.*[50] performed a case cohort study of postmenopausal women who did not have T2D and were enrolled in the Women's Health Initiative Observational Study, and they reported that fasting insulin levels were strongly positively associated with breast cancer risk independently of obesity. The authors concluded that hyperinsulinemia is an independent risk factor for postmenopausal breast cancer. Although this study was conducted with nondiabetic women, studies involving T2D women have also demonstrated an association between breast cancer risk and hyperinsulinemia. For example, in the Nurses' Health Study, women with T2D had a slightly elevated incidence of breast cancer, with a hazard ratio of 1.17, which was independent of body adiposity.[51] An increased risk of breast cancer was seen only in postmenopausal women and was predominantly seen in women with estrogen receptor–positive breast cancer.[51] In a meta-analysis performed by Larson and colleagues,[43] T2D patients were determined to have a 20% increased risk of breast cancer in comparison with nondiabetic subjects.[43]

Several studies have explored the relationship between T2D and breast cancer mortality and have shown a positive association. For example, Yancik *et al.*[52] reported that diabetic breast cancer patients had an increased risk of dying from breast cancer (relative risk = 1.76). Verlato and colleagues[53] studied over 3000 diabetic women and determined that their hazard ratio of dying from breast cancer was 1.40 in comparison with the general population. Coughlin *et al.*[54] performed a prospective cohort study with 590,000 patients and found that diabetic patients over the age of 30 years (the mean age of the participants at enrollment was 57 years) had an increased risk of breast cancer mortality in comparison with controls.

In addition to hyperinsulinemia, breast cancer and T2D are also thought to be related through the dysregulation of sex hormones. Estrogens may play a causal role in breast cancer.[55,56] Insulin can increase bioavailable estrogen by stimulating aromatase activity[57] and by inhibiting the production of sex hormone-binding globulin (SHBG), which leads to increased free estrogens.[58] SHBG levels have been inversely correlated with breast cancer risk in postmenopausal women.[59–62] Prospective studies performed on postmenopausal women showed that the women in the highest quintile of free estradiol and non–SHBG-bound estradiol levels had a relative risk of breast cancer of 2.58.[60] Postmenopausal obese women have increased androgen and estradiol production and decreased production of SHBG by the liver. This combination is associated with an increased risk of breast cancer.[63] Females with T2D have similar dysregulation of sex hormones and SHBG, which is likely mediated through chronic hyperinsulinemia.[64]

Taken together, several studies have demonstrated an increased risk of breast cancer and breast cancer–related mortality in patients with T2D. The increased breast cancer incidence and mortality in patients with T2D may be due to biological effects of the diabetic state, including hyperinsulinemia, activation of IR/IGF-R, and dysregulation of sex hormones.

Colorectal Cancer

Colorectal cancer is the second leading cause of cancer-related deaths in the United States. The American Cancer Society estimates that there will be 106,100 new cases of colon cancer and 40,870 new cases of rectal cancer in the United States in 2009. The disease is expected to cause about 49,920 deaths in 2009.[41] T2D has been suggested as a risk factor for colorectal cancer.[65] Some of the proposed mechanisms for the association between colorectal cancer and T2D are a slower bowel transit time among T2D patients with resultant increased exposure to toxins, increased production of carcinogenic bile acids, and hyperinsulinemia.[66] Of these potential mechanisms, hyperinsulinemia is the most thoroughly explored and appears to be a major player in the T2D-mediated increase in colorectal cancer risk.[67]

In vivo models have demonstrated a connection between insulin resistance and colorectal cancer. Insulin injected subcutaneously into Fisher 344 rats that had been injected with azoxymethane, a carcinogen, led to increased tumor incidence and burden in comparison with rats who did not receive insulin.[68] In another study, azoxymethane-treated Fisher 344 rats with high-fat diet–induced insulin resistance developed an increased number of aberrant crypt foci (precursors of colon cancer) in comparison with mice fed a control diet. In addition, there was a close correlation between the development of insulin resistance and tumor promotion.[69] Our earlier studies noted increased growth of mouse colon 38-adenocarcinoma cells in a model of diet-induced obesity and insulin resistance in comparison with control mice fed a regular chow diet.[70]

Studies in humans have also demonstrated a relationship between T2D and colorectal cancer. A prospective cohort of women between 30 and 55 years of age within the Nurses' Health Study demonstrated that T2D patients had a relative risk of 1.43 for colorectal cancer and a relative risk of 2.39 for fatal colorectal cancer.[71] A case control study of over 10,000 patients with colorectal cancer found that a diagnosis of T2D was associated with an increased risk of colorectal cancer (odds ratio = 1.42).[72] Larsson *et al.*[65] performed a meta-analysis of 15 studies and concluded that diabetes was associated with an increased risk of colorectal cancer, with an overall relative risk of 1.30. It is important to note that these authors were not able to distinguish between T1D and T2D in all of the patients studied in the analysis.

Several studies have shown an increased risk of colorectal cancer in patients with T2D on insulin therapy. In a retrospective cohort study of patients with T2D, the incidence of colorectal cancer in patients on insulin was higher than that in nonusers of insulin. The hazard ratio associated with more than 1 year of insulin therapy was 2.1. The odds ratio with each incremental year of insulin was 1.21.[73] A case control study demonstrated an increase in rectal cancer prevalence among Caucasian insulin users.[74]

Besides an increased risk of developing colon cancer, mortality from the disease is also affected by T2D, as shown by Coughlin and colleagues,[54] who reported increased colorectal cancer mortality in diabetic patients over the age of 30 years. Larsson *et al.*[65] also concluded that diabetes was positively related to colorectal cancer–related mortality.

Altogether, both colorectal cancer risk and mortality appear to be elevated in patients with T2D, and this association is likely to be mediated through hyperinsulinemia.

Pancreatic Cancer

In 2007, approximately 37,000 people were diagnosed with pancreatic cancer.[75] Pancreatic cancer was estimated to be responsible for 3% of cancer cases in men and 6% of cancer deaths in both genders in the United States in 2008.[41]

In vivo studies have shown a link between pancreatic cancer and T2D and obesity. Sprague-Dawley rats injected with 7,12-dimethylbenz[α]anthracene, a carcinogen, had an increased incidence of 7,12-dimethylbenz[α]anthracene–induced pancreatic cancer and dysplasia when they were fed a high-fat diet to induce obesity and insulin resistance.[76]

The stated prevalence of diabetes in pancreatic cancer has varied between 4% and 64%.[77] Pannala *et al.*[77] studied 512 patients with newly diagnosed pancreatic cancer and determined that T2D was 40% more prevalent in patients with pancreatic cancer in comparison with controls. An increased risk of pancreatic cancer in T2D patients has been widely reported.[54,78,79] In a meta-analysis by Huxley *et al.*,[80] an 80% increased risk of pancreatic cancer was seen in T2D patients. Ogunleye and colleagues[61] reported a 3-fold increased risk of pancreatic cancer in patients with T2D.

Diabetes may be both a risk factor for pancreatic cancer and a result of pancreatic cancer.[82] Diabetes may lead to pancreatic cancer through hyperinsulinemia.[82] Insulin has been shown to promote growth in human pancreatic cell lines.[83] Peripheral insulin resistance may enhance pancreatic carcinogenesis through increased cell turnover and enhanced proliferation of islet cells.[82] In T2D, exocrine pancreatic cells are subjected to high concentrations of insulin via the common blood supply between endocrine and exocrine pancreatic cells, and insulin acts as a mitogen on the exocrine cells leading to tumor growth.[84] However, this explanation does not explain the increased risk of pancreatic cancer observed in studies of T1D patients and insulin-treated diabetic patients.[85] On the other hand, diabetes may be a result of pancreatic cancer rather than a cause. Pancreatic cancer can lead to diabetes through islet cell destruction and through insulin resistance.[86,87] It is not yet clear exactly how pancreatic cancer can cause insulin resistance; however, it has been shown that among patients with pancreatic cancer, those with diabetes have elevated plasma levels of islet amyloid polypeptide, a hormonal factor secreted by pancreatic β cells that may cause insulin resistance.[86] Insulin resistance can occur early during the disease course of pancreatic cancer, and patients may present with diabetes long before developing signs of pancreatic cancer.[88] In fact, hyperglycemia may precede the diagnosis of pancreatic cancer by 24 months.[89] This concept of reverse causality was introduced by Gullo *et al.*,[88] who observed a relationship between diabetes and pancreatic cancer. However, no relationship was observed between diabetes and pancreatic cancer among long-term T2D patients. The increased incidence of diabetes in patients with pancreatic cancer was primarily due to recent-onset diabetes.[88] Therefore, Gullo and colleagues concluded that diabetes in these patients was likely due to their pancreatic tumors. In another study, patients with a recent diagnosis of diabetes had a 50% greater risk of malignancy than patients who had diabetes for at least 5 years.[80] In the aforementioned study by Pannala *et al.*,[77] 74% of diabetic patients with pancreatic cancer had recent-onset diabetes (defined as less than 2 years).[77]

Other studies have not corroborated this concept of reverse causality. Ogunleye and colleagues[81] noted an increased risk of pancreatic cancer in patients even after excluding patients diagnosed with diabetes in the year preceding their pancreatic cancer diagnosis. This observation suggests that T2D may contribute to pancreatic cancer. Thus, further studies are necessary to delineate the complicated relationship between diabetes and pancreatic cancer.

Prostate Cancer

Prostate cancer was estimated to be the cause of 25% of cancer cases and 10% of US cancer deaths in men in 2008.[41] Unlike the other cancers discussed thus far, prostate cancer has been shown to be less common in T2D patients. For example, Bonovas and colleagues[90] performed a meta-analysis of 14 studies published through 2002 and reported a 9% decreased risk of prostate carcinoma in diabetic patients. Kasper and Giovannucci[91] expanded on the work of Bonovas *et al.* by performing a meta-analysis that included more than twice the number of patients included in the initial analysis. They reported that diabetic men had a 16% decreased risk of developing prostate carcinoma. However, both of these meta-analyses included studies that did not distinguish between T1D and T2D.

High testosterone levels have been associated with an increased risk of prostate cancer.[91,92] T2D men commonly have low testosterone levels, and this may explain the negative association between T2D and prostate cancer. Many men with T2D are also obese and elderly, and both factors are associated with decreased testosterone levels. Proposed mechanisms for the androgen deficiency seen in T2D and obese men include the following: lower SHBG, suppression of gonadotropin release or Leydig cell testosterone production, increased levels of inflammatory cytokines that inhibit steroidogenesis, and increased aromatase activity resulting in increased conversion of testosterone to estradiol.[93] It is not yet clear if T2D has an independent effect on testosterone levels or if the effect is mediated by obesity. However, some studies have suggested that the link between prostate cancer and T2D is through the effects of hyperinsulinemia on testosterone levels. Studies in healthy men have shown an inverse correlation between total testosterone levels and insulin levels.[94] A study of 574 T2D men found that 43% had reduced total testosterone levels and 57% had decreased calculated free testosterone levels. Reduced free testosterone levels were independently correlated with insulin resistance.[95]

Several studies have shown a negative association between hyperinsulinemia and hyperglycemia and prostate cancer risk. Stocks and colleagues[96] performed a prospective study and reported that increased insulin resistance and worsened glycemic control decreased prostate cancer risk. Baradaran *et al.*[97] described an inverse relationship between diabetes duration and prostate cancer risk.

Although prostate cancer risk appears to be decreased in diabetic patients, patients with diabetes and insulin resistance who develop prostate cancer may have a higher risk of cancer-related mortality than nondiabetic subjects. Ma *et al.*[98] found that higher C-peptide concentrations were predictive of prostate cancer–related mortality and worse outcome. Higher insulin levels promoting tumor growth have been suggested as an explanation for this increase in mortality. On the other hand, Coughlin *et al.*[54] reported a lower risk of prostate cancer mortality in men with T2D. Thus, the relationship between prostate cancer mortality and T2D is not yet clear. In the other specific cancer types reviewed previously, diabetes has been associated with an increased risk of cancer and cancer-related mortality; however, in prostate cancer, diabetes appears to be associated with a decreased risk of cancer. There are conflicting data on the relationship between T2D and prostate cancer mortality, so further study is necessary.

Taken together, several studies have shown a lower risk of prostate cancer in patients with T2D. Proposed mechanisms for this decrease have centered on the effects of hyperinsulinemia on testosterone levels. The lower risk of prostate cancer in T2D is contrary to the positive associations between T2D and the other cancers discussed previously. It is possible that the beneficial effects of low testosterone outweigh the negative effects of hyperinsulinemia. Given the noted discrepancies between prostate cancer and the other cancers reviewed, more studies are necessary to further define the relationship between prostate cancer and T2D.

Associations between T2D and different types of cancer have been proposed, and hyperinsulinemia appears to play a major role in these associations. T2D has been clearly linked to breast cancer and colorectal cancer. Although a positive relationship between T2D and pancreatic cancer has been demonstrated, it is unclear if T2D is a cause or result of pancreatic cancer. On the other hand, T2D has been shown to be inversely related to prostate cancer. This is most likely related to lower testosterone levels found in T2D men as a result of hyperinsulinemia. Although much work has been done to explore the association between T2D and specific cancers, further data will be helpful to better delineate these individual relationships and their underlying mechanisms.

T2D has been clearly linked to breast cancer and colorectal cancer. Although a positive relationship between T2D and pancreatic cancer has been demonstrated, it is unclear if T2D is a cause or result of pancreatic cancer. On the other hand, T2D has been shown to be inversely related to prostate cancer.

EFFECTS OF MEDICATIONS ON THE ASSOCIATION BETWEEN TYPE 2 DIABETES MELLITUS AND CANCER

T2D is treated with a variety of medications, so there may also be a link between these medications and cancer. The therapies employed in the treatment of T2D include medications that increase insulin in the circulation (eg, sulfonylureas and insulin) and therapies that improve insulin action and thus decrease insulin levels [ie, metformin and thiazolidinediones (TZDs)]. Studies performed in humans have largely been conducted with diabetic patients as they are already taking these medications; therefore, it is not clear if the association between these medications and cancer is due to a direct effect of the medications on tumor promotion and growth or is due to indirect effects through the medications' effect on insulin levels. On the other hand, many of the in vivo studies have been performed with nondiabetic animals and have thus demonstrated a direct effect. However, as these in vivo studies did not determine insulin levels, an indirect mechanism cannot be completely ruled out. The following sections discuss the aforementioned medications and their connection to cancer in detail.

Sulfonylureas

Sulfonylureas are insulin secretagogues that act by increasing insulin release from the β cells of the pancreas. Sulfonylureas bind to the adenosine triphosphate–dependent potassium channel in pancreatic β cells. This leads to inhibition of these channels, which alters the resting potential of the cell and leads to calcium influx and stimulation of insulin secretion.[99] Potassium channels regulate the growth and proliferation of many cell types.[100] Thus, it has been suggested that sulfonylureas, which inhibit potassium channels, may have an antiproliferative effect on tumor cells. In vitro studies have demonstrated an antitumor effect of sulfonylureas. For example, the sulfonylurea glibenclamide inhibited the growth of HTB-9 human urinary bladder carcinoma cells in vitro.[101] Glibenclamide also inhibited cell proliferation of the human liver epithelial cancer cell lines HepG2, HuH-7, and HFL[100] and the human gastric cancer cell line MGC–803.[102] In vivo studies with sulfonylureas, though few in number, have also demonstrated an antitumor effect. For example, sulofenur and LY295 501, both diarylsulfonylureas, showed significant activity against colon tumor xenografts.[103] Treatment with the novel sulfonylurea DW2282 led to reduced tumor growth in nondiabetic mice transplanted with human tumor cell lines.[104] Thus, sulfonylureas appear to have an antiproliferative effect on cancer cells in vitro and in vivo. Human studies have not demonstrated this same antiproliferative effect. For example, Bowker *et al.*[105] found a 30% increase in the risk of cancer-related mortality among patients treated with sulfonylureas in comparison with patients on metformin. In a study of patients with T2D treated with sulfonylureas or insulin, the odds ratio of developing hepatocellular carcinoma was 2.99 versus 0.33 in patients treated with metformin.[106] Although this study did not separate out sulfonylurea and insulin use, the results do demonstrate a positive association between sulfonylurea use and cancer. The disparity between the results from in vitro and animal studies and human studies may be due to the indirect effects of sulfonylureas. In vivo studies have largely been performed in nondiabetic animals. Therefore, the direct antitumor properties of the sulfonylureas have been observed. On the other hand, for the most part, human studies have been conducted with diabetic patients, in whom both the direct and indirect effects of the sulfonylureas are at play. Thus, the indirect proliferative effect of elevated insulin levels may override the direct antitumor effects. However, given the limited number of studies, further data are needed to evaluate how sulfonylureas influence cancer and cancer-related mortality.

Insulin

Oral agents become less effective in the treatment of T2D as β cell function declines over the course of the disease. Patients with persistent hyperglycemia despite oral hypoglycemic therapy may need to have insulin added to their oral medication regimen. The mitogenic effects of human insulin have been noted in vitro, in vivo, and in human studies.[44–47,68] Bowker *et al.*[105] performed a retrospective cohort study of patients newly diagnosed with T2D and found a 90% relative increase in the risk of cancer-related mortality in patients treated with insulin versus patients on metformin. In a retrospective cohort study by Currie and colleagues,[107] patients treated with insulin had an increased risk of colorectal and pancreatic cancer but not of breast or prostate cancer in comparison with patients treated with metformin. Yang *et al.*[73] reported that the length of insulin treatment in patients with T2D was positively correlated with an increased incidence of colorectal cancer. Insulin analogues have been developed as an alternative to human insulin for the treatment of T1D and T2D. Insulin analogues are produced via recombinant DNA technology in which the insulin molecule is modified to make it more physiological; this can improve glycemic control in patients. Manipulation of the structure of the insulin molecule can lead to altered binding to its receptors. Concern has centered on the carcinogenicity of these modifications. Studies have shown mixed results. The insulin analogues B10Asp and glargine were found to have increased mitogenic effects on human osteosarcoma cells (Saos/B10) in vitro, whereas the insulin analogues lispro and aspart were slightly less mitogenic than human insulin.[108] Enhanced proliferation of MCF-7 cells was seen in cells treated with insulin glargine versus human insulin.[109] In vitro studies demonstrated that insulin glargine, insulin detemir, and insulin lispro led to increased cancer cell proliferation in HCT-116 colorectal cancer cells, PC-3 prostate cancer cells, and MCF-7 breast cancer cells in comparison with human insulin.[110] Stammberger and colleagues[111] studied the effects of insulin glargine on mammary tumors in rats and mice and found no difference between animals treated with insulin glargine and human insulin. Animal studies have been limited, however, by the excessive mortality encountered as a result of high-dose insulin glargine.[111] Currie *et al.* found no difference in the cancer risk in patients treated with insulin glargine and those treated with human insulin. Hemkens *et al.*[112] reported a dose-dependent cancer risk in patients treated with insulin glargine versus human insulin. A follow-up study did not find an increased risk of all cancers or site-specific cancers in patients treated with insulin glargine.[113] Given the conflicting results of existing studies, further studies are necessary to determine the mitogenic potential of the insulin analogues.

Insulin has been shown to have direct proliferative effects. The mitogenic potential of insulin analogues has not yet been conclusively determined, so further studies are necessary.

Insulin has been shown to have direct proliferative effects. The mitogenic potential of insulin analogues has not yet been conclusively determined, so further studies are necessary.

Metformin

Metformin is classified as a biguanide. It is commonly called an insulin sensitizer. Metformin acts by decreasing hepatic gluconeogenesis and thereby decreasing hepatic glucose production and release. In addition, metformin enhances glucose uptake in the muscles and thus improves insulin sensitivity. One mechanism of action of metformin involves the activation of adenosine monophosphate–dependent protein kinase (AMPK) through its upstream kinase, liver kinase B1, a tumor suppressor gene involved in epithelial neoplasia (Peutz-Jeghers syndrome).[114–116] AMPK appears to suppress cell proliferation via inhibition of the mammalian target of rapamycin pathway and through the up-regulation of the p53-p21 axis, reduction of cyclin D1 levels, and suppression of cyclin-dependent kinases leading to G1 cell cycle arrest.[117]

In vitro studies using MCF-7 human breast cancer cells have shown that metformin up-regulates AMPK activity and reduces general protein synthesis and thereby functions as a growth inhibitor.[118] Metformin also had an inhibitory effect on the growth of PC-3 prostate cancer cells and HT-29 colon cancer cells through AMPK activation.[119] Recently, Hirsch *et al.*[120] showed that metformin selectively killed cancer stem cells in 4 genetically different types of breast cancer cells.

Ben Sahra *et al.*[121] studied the effect of metformin on human prostate cancer cells in vitro and in vivo. Metformin was found to inhibit cell proliferation of malignant cells in a dose-dependent fashion by decreasing cyclin D1 protein levels and blocking the cell cycle. Female transgenic HER-2/neu mice chronically treated with metformin had decreased incidence and size of mammary adenocarcinomas. The mean latency of tumor development was increased with metformin treatment.[122] Metformin has also been shown to reduce tumor growth and increase apoptosis of p53-deficient colon cancer cell lines in vivo.[123] All the aforementioned studies were performed in nondiabetic mice and thus show that metformin has a direct antitumor effect in vitro and in vivo.

Using a subcutaneous injection of Lewis lung carcinoma 1 cells, Algire and colleagues[124] studied the effect of metformin on high-energy diet–induced tumor progression. Mice on the high-energy diet had a higher tumor volume than mice receiving the control diet. Metformin treatment attenuated the stimulatory effect of the high-energy diet on tumor growth but did not affect the tumor growth of mice on the control diet.[124] This suggests that both the insulin-reducing action of metformin and its antitumor activity may be involved in inhibiting tumor growth in this tumor model.

In human studies, metformin has also been shown to reduce the risk of cancer in patients with T2D. Evans *et al.*[125] showed that metformin treatment was associated with a decreased risk of cancer in patients with newly diagnosed T2D (odds ratio = 0.77). The authors studied the duration, total number of prescriptions dispensed, and cumulative dose of metformin. A dose-response relationship was seen, with lower rates of cancer in patients who had a higher degree of exposure to metformin. In the aforementioned study by Curie *et al.*,[107] patients with diabetes treated with metformin monotherapy had the lowest rates of cancer. There was an increased risk of cancer in patients treated with sulfonylureas and insulin. The addition of metformin to these medications decreased most of this excess risk. Jiralerspong *et al.*[126] studied 2529 patients receiving neoadjuvant chemotherapy for early-stage breast cancer and observed their rates of pathological complete response, which was defined as the absence of tumor in tissue removed at the time of surgery. Diabetic patients receiving metformin had a pathological complete response rate of 24% versus a pathological complete response rate of 8% among diabetics not receiving metformin.

Metformin appears to have an antineoplastic effect in vitro, in vivo, and in human subjects.

Metformin appears to have an antineoplastic effect in vitro, in vivo, and in human subjects.

The antiproliferative effect is likely due to both direct and indirect mechanisms. Direct effects are thought to be related to AMPK activation leading to tumor suppression. Metformin also decreases insulin levels and thus indirectly inhibits cancer cell growth.

Thiazolidinediones

Peroxisome proliferator-activator receptors (PPARs) are members of the nuclear hormone superfamily and are important in the regulation of lipid homeostasis. Three distinct PPARs–PPAR-α, PPAR-δ, and PPAR-γ –exist, and PPAR-γ is the best defined.[127] PPAR-γ is a transcription factor involved in the differentiation of cells such as adipocytes and macrophage foam cells that have active lipid metabolism. TZDs, a class of commonly prescribed antidiabetic medications, include pioglitazone, rosiglitazone, ciglitazone, and troglitazone (removed from the market because of hepatotoxicity). The mechanism of action of TZDs involves the activation of PPAR-γ, which results in an improvement in the insulin sensitivity in peripheral tissue (mainly the adipose tissue, liver, and skeletal muscles). As PPAR-γ is involved in differentiation, it might serve as a target for antineoplastic therapy. However, studies on the antitumorigenic potential of TZDs have yielded conflicting results.

PPAR-γ has been noted to be overexpressed in liposarcomas and colon, breast, and prostate carcinomas.[128–131] Overexpression has been observed in both well and poorly differentiated tumors. In vitro studies on human liposarcoma cells have shown that after treatment with the PPAR-γ ligand pioglitazone, these cancer cells undergo terminal differentiation.[128] Troglitazone inhibited growth and induced differentiation of colon cancer cells in vitro.[131] In vitro studies of lung, pancreatic, and hematopoietic tumor cells also demonstrated decreased growth with TZD treatment.[130–133] TZD treatment decreased tumor volume and weight in nondiabetic mice injected with MCF-7 cancer cells.[133] In vivo studies of colon cancer cells and troglitazone showed that treatment with troglitazone led to a reduced tumor growth rate in nondiabetic mice.[131]

Other studies have reported treatment with PPAR-γ agonists to be tumor-promoting. In nondiabetic C57BL/6J-APC$^{Min/+}$ mice predisposed to intestinal neoplasia, treatment with PPAR-γ agonists (troglitazone and rosiglitazone) led to increased frequency and size of colon tumors.[134] Saez and colleagues[135] studied nondiabetic mice with a predilection to intestinal neoplasia and found that mice treated with troglitazone developed considerably more polyps. This same group later generated mice with a constitutively active form of PPAR-γ in the mammary gland. Constitutive activation of PPAR-γ had no effect on mammary gland differentiation or function. However, in mice with increased proclivity to breast cancer, there was acceleration of tumor formation and mortality.[127] A factor confounding the comparison between these studies is that the studies showing tumor inhibition were primarily performed with in vitro models, whereas the studies showing tumor promotion used in vivo models, and this suggests that the local environment of the tumor may account for the differing effects of PPAR-γ stimulation.

Studies of the effects of TZDs on cancer in humans have also been inconsistent. The Prospective Pioglitazone Clinical Trial in Macrovascular Events (PROACTIVE), which studied over 5000 diabetic patients treated with pioglitazone or placebo, reported a reduction in the incidence of breast carcinoma in patients treated with pioglitazone; however, this reduction was not significant.[136] Govindarajan and colleagues[137] performed a retrospective analysis of lung cancer patients and found that lung cancer risk was reduced by 33% in patients taking TZDs. In the same study by Govindarajan and colleagues, the risk reduction for colorectal and prostate cancers in patients on TZDs was not statistically significant. Troglitazone did not show any benefit in 22 metastatic breast cancer patients; however, this study was stopped when troglitazone was taken off the market.[138]

Given the conflicting results of existing studies, further studies are needed to better define the role of PPAR-γ activation and TZDs in cancer. In addition, TZDs may have effects on cancer development and growth that are independent of PPAR-γ,[139] and these must be elucidated before any conclusions on TZD use in cancer can be made.

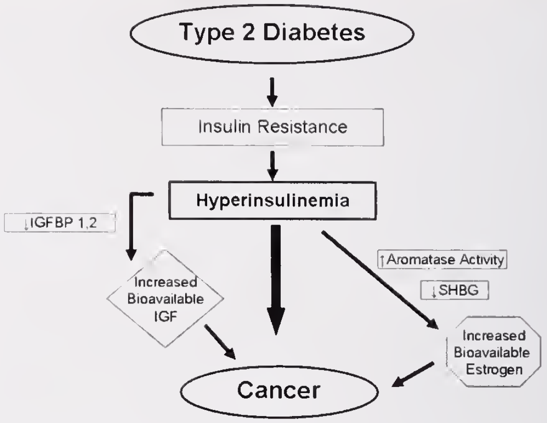


Fig 3. Potential mechanism for the link between type 2 diabetes and cancer. In patients with type 2 diabetes, insulin resistance results in elevated insulin levels. Hyperinsulinemia may lead to cancer through its direct effects on cancer cells via IR and/or IGF-IR. Alternatively, insulin may act indirectly on cancer cells. Insulin increases bioavailable estrogen by stimulating aromatase activity and by inhibiting the production of SHBG. In addition, insulin increases the amount of bioavailable IGF as it inhibits the production of 2 of the IGFBPs, IGFBP-1 and IGFBP-2, which transport IGF-I to the peripheral tissues. **Abbreviations:** IGF, insulin-like growth factor; IGF-IR, insulin-like growth factor I receptor; IGFBP, insulin-like growth factor binding protein; IR, insulin receptor; SHBG, sex hormone-binding globulin.

The medications used to treat T2D include therapies that increase insulin in the circulation and those that enhance sensitivity to insulin. Given the role that hyperinsulinemia is thought to play in the development of cancer in T2D patients, therapies that decrease insulin resistance would be presumed to lower cancer risk more than those that increase circulating insulin. The relationship between T2D medications and cancer may be due to a direct effect of the medication or an indirect mechanism through effects on insulin levels. Sulfonylureas appear to have a direct antitumor effect in animal models that is overcome by their indirect effects on insulin levels in human subjects.

Sulfonylureas appear to have a direct antitumor effect in animal models that is overcome by their indirect effects on insulin levels in human subjects.

Insulin has a direct mitogenic effect, and it is not yet clear if the insulin analogues are more mitogenic than human insulin. Metformin appears to have a direct effect as animal studies in nondiabetic mice have demonstrated an anticancer effect of the medication. TZDs appear to have a direct effect as well, although it is not yet apparent if this effect is tumor-promoting or tumor-suppressing. Data are limited on all drugs, so further study of all of these medications and their effects on cancer risk and mortality are warranted.

CONCLUSION

Epidemiological studies have shown a link between T2D and cancer. Positive correlations between T2D and cancers of the breast, colon, and pancreas have been observed in vitro, in animal studies, and in human studies. Hyperinsulinemia, through its effects on the insulin and IGF pathways, is thought to play a major role in this association (Figure 3). The effects of insulin and IGF-1 on cancer development and progression have been demonstrated in animal and human studies. Insulin resistance and the resulting hyperinsulinemia may serve as a potential target for cancer therapy. In addition, early detection and treatment of hyperinsulinemia, which often occurs many years before overt T2D develops, may decrease the cancer burden and may be an important component of cancer prevention and care.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Centers for Disease Control and Prevention. 2007; national diabetes fact sheet. http://www.cdc.gov/diabetes/pubs/estimates07.htm. Accessed January 2010.
2. Kaaks R, Lukanova A. Energy balance and cancer: the role of insulin and insulin-like growth factor-I. *Proc Nutr Soc* 2001; 60: 91–106.
3. Polednak AP. Trends in incidence rates for obesity-associated cancers in the US. *Cancer Detect Prev* 2003; 27: 415–421.
4. Calle EE, Kaaks R. Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. *Nat Rev Cancer* 2004; 4: 579–591.
5. Xue F, Michels KB. Diabetes, metabolic syndrome, and breast cancer: a review of the current evidence. *Am J Clin Nutr* 2007; 86: s823–s835.
6. Vona-Davis L, Howard-McNatt M, Rose DP. Adiposity, type 2 diabetes and the metabolic syndrome in breast cancer. *Obes Rev* 2007; 8: 395–408.
7. Zendehdel K, Nyren O, Ostenson CG, et al. Cancer incidence in patients with type 1 diabetes mellitus: a population-based cohort study in Sweden. *J Natl Cancer Inst* 2003; 95: 1797–1800.
8. LeRoith D, Werner H, Beitner-Johnson D, Roberts CT Jr. Molecular and cellular aspects of the insulin-like growth factor I receptor. *Endocr Rev* 1995; 16: 143–163.
9. Samani AA, Yakar S, LeRoith D, Brodt P. The role of the IGF system in cancer growth and metastasis: overview and recent insights. *Endocr Rev* 2007; 28: 20–47.
10. Taha C, Klip A. The insulin signaling pathway. *J Membr Biol* 1999; 169: 1–12.
11. Frasca F, Pandini G, Sciacca L, et al. The role of insulin receptors and IGF-I receptors in cancer and other diseases. *Arch Physiol Biochem* 2008; 114: 23–37.
12. Saltiel AR, Kahn CR. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* 2001; 414: 799–806.
13. Fresno Vara JA, Casado E, de Castro J, et al. PI3K/Akt signalling pathway and cancer. *Cancer Treat Rev* 2004; 30: 193–204.
14. Wolf I, Seger R. The mitogen-activated protein kinase signaling cascade: from bench to bedside. *Isr Med Assoc J* 2002; 4: 641–647.
15. Belfiore A. The role of insulin receptor isoforms and hybrid insulin/IGF-I receptors in human cancer. *Curr Pharm Des* 2007; 13: 671–686.
16. Mosthaf L, Grako K, Dull TJ, et al. Functionally distinct insulin receptors generated by tissue-specific alternative splicing. *EMBO J* 1990; 9: 2409–2413.
17. Sciacca L, Prisco M, Wu A, et al. Signaling differences from the A and B isoforms of the insulin receptor (IR) in 32D cells in the presence or absence of IR substrate-1. *Endocrinology* 2003; 144: 2650–2658.
18. Frasca F, Pandini G, Scalia P, et al. Insulin receptor isoform A, a newly recognized, high-affinity insulin-like growth factor II receptor in fetal and cancer cells. *Mol Cell Biol* 1999; 19: 3278–3288.
19. Sciacca L, Costantino A, Pandini G, et al. Insulin receptor activation by IGF-II in breast cancers: evidence for a new autocrine/paracrine mechanism. *Oncogene* 1999; 18: 2471–2479.
20. Pandini G, Frasca F, Mineo R, et al. Insulin/insulin-like growth factor I hybrid receptors have different biological characteristics depending on the insulin receptor isoform involved. *J Biol Chem* 2002; 277: 39684–39695.
21. Pandini G, Vigneri R, Costantino A, et al. Insulin and insulin-like growth factor-I (IGF-I) receptor overexpression in breast cancers leads to insulin/IGF-I hybrid receptor overexpression: evidence for a second mechanism of IGF-I signaling. *Clin Cancer Res* 1999; 5: 1935–1944.
22. Soos MA, Field CE, Siddle K. Purified hybrid insulin/insulin-like growth factor-I receptors bind insulin-like growth factor-I, but not insulin, with high affinity. *Biochem J* 1993; 290(pt 2): 419–426.
23. Bailyes EM, Nave BT, Soos MA, et al. Insulin receptor/IGF-I receptor hybrids are widely distributed in mammalian tissues: quantification of individual receptor species by selective immunoprecipitation and immunoblotting. *Biochem J* 1997; 327(pt 1): 209–215.
24. Papa V, Pezzino V, Costantino A, et al. Elevated insulin receptor content in human breast cancer. *J Clin Invest* 1990; 86: 1503–1510.
25. Papa V, Gliozzo B, Clark GM, et al. Insulin-like growth factor-I receptors are overexpressed and predict a low risk in human breast cancer. *Cancer Res* 1993; 53: 3736–3740.
26. Frittitta L, Sciacca L, Catalfamo R, et al. Functional insulin receptors are overexpressed in thyroid tumors: is this an early event in thyroid tumorigenesis? *Cancer* 1999; 85: 492–498.
27. Vella V, Pandini G, Sciacca L, et al. A novel autocrine loop involving IGF-II and the insulin receptor isoform-A stimulates growth of thyroid cancer. *J Clin Endocrinol Metab* 2002; 87: 245–254.

28. Kuramoto H, Hongo A, Liu YX, et al. Immunohistochemical evaluation of insulin-like growth factor I receptor status in cervical cancer specimens. *Acta Med Okayama* 2008; 62: 251–259.
29. Railo MJ, von Smitten K, Pekonen F. The prognostic value of insulin-like growth factor-I in breast cancer patients. Results of a follow-up study on 126 patients. *Eur J Cancer* 1994; 30A: 307–311.
30. Matsubara J, Yamada Y, Hirashima Y, et al. Impact of insulin-like growth factor type 1 receptor, epidermal growth factor receptor, and HER2 expressions on outcomes of patients with gastric cancer. *Clin Cancer Res* 2008; 14: 3022–3029.
31. Bentov I, Narla G, Schayek H, et al. Insulin-like growth factor-I regulates Kruppel-like factor-6 gene expression in a p53-dependent manner. *Endocrinology* 2008; 149: 1890–1897.
32. Bruchim I, Attias Z, Werner H. Targeting the IGF1 axis in cancer proliferation. *Expert Opin Ther Targets* 2009; 13: 1179–1192.
33. Sarfstein R, Werner H. The WT1 Wilms' tumor suppressor gene is a downstream target for insulin-like growth factor-I (IGF-I) action in PC12 cells. *J Neurochem* 2006; 99: 818–826.
34. Webster NJ, Resnik JL, Reichart DB, et al. Repression of the insulin receptor promoter by the tumor suppressor gene product p53: a possible mechanism for receptor overexpression in breast cancer. *Cancer Res* 1996; 56: 2781–2788.
35. Frittitta L, Cerrato A, Sacco MG, et al. The insulin receptor content is increased in breast cancers initiated by three different oncogenes in transgenic mice. *Breast Cancer Res Treat* 1997; 45: 141–147.
36. Mathieu MC, Clark GM, Allred DC, et al. Insulin receptor expression and clinical outcome in node-negative breast cancer. *Proc Assoc Am Physicians* 1997; 109: 565–571.
37. Law JH, Habibi G, Hu K, et al. Phosphorylated insulin-like growth factor-I/insulin receptor is present in all breast cancer subtypes and is related to poor survival. *Cancer Res* 2008; 68: 10238–10246.
38. Sciacca L, Mineo R, Pandini G, et al. In IGF-I receptor-deficient leiomyosarcoma cells autocrine IGF-II induces cell invasion and protection from apoptosis via the insulin receptor isoform A. *Oncogene* 2002; 21: 8240–8250.
39. Belfiore A, Pandini G, Vella V, et al. Insulin/IGF-I hybrid receptors play a major role in IGF-I signaling in thyroid cancer. *Biochimie* 1999; 81: 403–407.
40. Vigneri P, Frasca F, Sciacca L, et al. Diabetes and cancer. *Endocr Relat Cancer* 2009; 16: 1103–1123.
41. *2008 American Cancer Society Facts and Figures.* Atlanta, GA: American Cancer Society; 2008.
42. Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. *Lancet Oncol* 2001; 2: 133–140.
43. Larsson SC, Mantzoros CS, Wolk A. Diabetes mellitus and risk of breast cancer: a meta-analysis. *Int J Cancer* 2007; 121: 856–862.
44. Ish-Shalom D, Christoffersen CT, Vorwerk P, et al. Mitogenic properties of insulin and insulin analogues mediated by the insulin receptor. *Diabetologia* 1997; 40(suppl 2): S25–S31.
45. Chappell J, Leitner JW, Solomon S, et al. Effect of insulin on cell cycle progression in MCF-7 breast cancer cells. Direct and potentiating influence. *J Biol Chem* 2001; 276: 38023–38028.
46. Shafie SM, Hilf R. Insulin receptor levels and magnitude of insulin-induced responses in 7,12-dimethylbenz(a)anthracene-induced mammary tumors in rats. *Cancer Res* 1981; 41: 826–829.
47. Heuson JC, Legros N, Heimann R. Influence of insulin administration on growth of the 7,12-dimethylbenz(a)anthracene-induced mammary carcinoma in intact, oophorectomized, and hypophysectomized rats. *Cancer Res* 1972; 32: 233–238.
48. Nunez NP, Oh WJ, Rozenberg J, et al. Accelerated tumor formation in a fatless mouse with type 2 diabetes and inflammation. *Cancer Res* 2006; 66: 5469–5476.
49. Novosyadlyy R, Lann DE, Vijayakumar A, et al. Insulin-mediated acceleration of breast cancer development and progression in a non-obese model of type 2 diabetes. *Cancer Res* 2010; 70: 741.
50. Gunter MJ, Hoover DR, Yu H, et al. Insulin, insulin-like growth factor-I, and risk of breast cancer in postmenopausal women. *J Natl Cancer Inst* 2009; 101: 48–60.
51. Michels KB, Solomon CG, Hu FB, et al. Type 2 diabetes and subsequent incidence of breast cancer in the Nurses' Health Study. *Diabetes Care* 2003; 26: 1752–1758.
52. Yancik R, Wesley MN, Ries LA, et al. Effect of age and comorbidity in postmenopausal breast cancer patients aged 55 years and older. *JAMA* 2001; 285: 885–892.
53. Verlato G, Zoppini G, Bonora E, Muggeo M. Mortality from site-specific malignancies in type 2 diabetic patients from Verona. *Diabetes Care* 2003; 26: 1047–1051.
54. Coughlin SS, Calle EE, Teras LR, et al. Diabetes mellitus as a predictor of cancer mortality in a large cohort of US adults. *Am J Epidemiol* 2004; 159: 1160–1167.
55. Berrino F, Muti P, Micheli A, et al. Serum sex hormone levels after menopause and subsequent breast cancer. *J Natl Cancer Inst* 1996; 88: 291–296.
56. Cauley JA, Lucas FL, Kuller LH, et al. Elevated serum estradiol and testosterone concentrations are associated with a high risk for breast cancer. Study of Osteoporotic Fractures Research Group. *Ann Intern Med* 1999; 130: 270–277.
57. McTernan PG, Anwar A, Eggo MC, et al. Gender differences in the regulation of P450 aromatase expression and activity in human adipose tissue. *Int J Obes Relat Metab Disord* 2000; 24: 875–881.
58. Singh A, Hamilton-Fairley D, Koistinen R, et al. Effect of insulin-like growth factor-type I (IGF-I) and insulin on the secretion of sex hormone binding globulin and IGF-I binding protein (IBP-I) by human hepatoma cells. *J Endocrinol* 1990; 124: R1–R3.
59. Toniolo PG, Levitz M, Zeleniuch-Jacquotte A, et al. A prospective study of endogenous estrogens and breast cancer in postmenopausal women. *J Natl Cancer Inst* 1995; 87: 190–197.
60. Key T, Appleby P, Barnes I, Reeves G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst* 2002; 94: 606–616.
61. Zeleniuch-Jacquotte A, Shore RE, Koenig KL, et al. Postmenopausal levels of oestrogen, androgen, and SHBG and breast cancer: long-term results of a prospective study. *Br J Cancer* 2004; 90: 153–159.

62. Kaaks R, Rinaldi S, Key TJ, et al. Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. *Endocr Relat Cancer* 2005; 12: 1071–1082.
63. Endogenous Hormones and Breast Cancer Collaborative Group. Free estradiol and breast cancer risk in postmenopausal women: comparison of measured and calculated values. *Cancer Epidemiol Biomarkers Prev* 2003; 12: 1457–1461.
64. Haffner SM, Laakso M, Miettinen H, et al. Low levels of sex hormone-binding globulin and testosterone are associated with smaller, denser low density lipoprotein in normoglycemic men. *J Clin Endocrinol Metab* 1996; 81: 3697–3701.
65. Larsson SC, Orsini N, Wolk A. Diabetes mellitus and risk of colorectal cancer: a meta-analysis. *J Natl Cancer Inst* 2005; 97: 1679–1687.
66. Will JC, Galuska DA, Vinicor F, Calle EE. Colorectal cancer: another complication of diabetes mellitus? *Am J Epidemiol* 1998; 147: 816–825.
67. Giovannucci E. Insulin and colon cancer. *Cancer Causes Control* 1995; 6: 164–179.
68. Tran TT, Medline A, Bruce WR. Insulin promotion of colon tumors in rats. *Cancer Epidemiol Biomarkers Prev* 1996; 5: 1013–1015.
69. Koohestani N. Chia MC, Pham NA, et al. Aberrant crypt focus promotion and glucose intolerance: correlation in the rat across diets differing in fat, n-3 fatty acids and energy. *Carcinogenesis* 1998; 19: 1679–1684.
70. Yakar S, Nunez NP, Pennisi P, et al. Increased tumor growth in mice with diet-induced obesity: impact of ovarian hormones. *Endocrinology* 2006; 147: 5826–5834.
71. Hu FB, Manson JE, Liu S, et al. Prospective study of adult onset diabetes mellitus (type 2) and risk of colorectal cancer in women. *J Natl Cancer Inst* 1999; 91: 542–547.
72. Yang YX, Hennessy S, Lewis JD. Type 2 diabetes mellitus and the risk of colorectal cancer. *Clin Gastroenterol Hepatol* 2005; 3: 587–594.
73. Yang YX, Hennessy S, Lewis JD. Insulin therapy and colorectal cancer risk among type 2 diabetes mellitus patients. *Gastroenterology* 2004; 127: 1044–1050.
74. Vinikoor LC, Long MD, Keku TO, et al. The association between diabetes, insulin use, and colorectal cancer among whites and African Americans. *Cancer Epidemiol Biomarkers Prev* 2009; 18: 1239–1242.
75. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2009. *CA Cancer J Clin* 2009; 59: 225–249.
76. Z'Graggen K, Warshaw AL, Werner J, et al. Promoting effect of a high-fat/high-protein diet in DMBA-induced ductal pancreatic cancer in rats. *Ann Surg* 2001; 233: 688–695.
77. Pannala R, Leirness JB, Bamlet WR, et al. Prevalence and clinical profile of pancreatic cancer-associated diabetes mellitus. *Gastroenterology* 2008; 134: 981–987.
78. Rousseau MC, Parent ME, Pollak MN, Siemiatycki J. Diabetes mellitus and cancer risk in a population-based case-control study among men from Montreal, Canada. *Int J Cancer* 2006; 118: 2105–2109.
79. Kuriki K, Hirose K, Tajima K. Diabetes and cancer risk for all and specific sites among Japanese men and women. *Eur J Cancer Prev* 2007; 16: 83–89.
80. Huxley R, Ansary-Moghaddam A, Berrington de Gonzalez A, et al. Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. *Br J Cancer* 2005; 92: 2076–2083.
81. Ogunleye AA, Ogston SA, Morris AD, Evans JM. A cohort study of the risk of cancer associated with type 2 diabetes. *Br J Cancer* 2009; 101: 1199–1201.
82. Silverman DT, Schiffman M, Everhart J, et al. Diabetes mellitus, other medical conditions and familial history of cancer as risk factors for pancreatic cancer. *Br J Cancer* 1999; 80: 1830–1837.
83. Fisher WE, Boros LG, Schirmer WJ. Insulin promotes pancreatic cancer: evidence for endocrine influence on exocrine pancreatic tumors. *J Surg Res* 1996; 63: 310–313.
84. Williams JA, Goldfine ID. The insulin-pancreatic acinar axis. *Diabetes* 1985; 34: 980–986.
85. Stevens RJ, Roddam AW, Beral V. Pancreatic cancer in type 1 and young-onset diabetes: systematic review and meta-analysis. *Br J Cancer* 2007: 96: 507–509.
86. Permert J, Larsson J, Westermark GT, et al. Islet amyloid polypeptide in patients with pancreatic cancer and diabetes. *N Engl J Med* 1994; 330: 313–318.
87. Isaksson B, Strommer L, Friess H, et al. Impaired insulin action on phosphatidylinositol 3-kinase activity and glucose transport in skeletal muscle of pancreatic cancer patients. *Pancreas* 2003; 26: 173–177.
88. Gullo L, Pezzilli R, Morselli-Labate AM. Diabetes and the risk of pancreatic cancer. *N Engl J Med* 1994; 331: 81–84.
89. Pannala R, Leibson CL, Rabe KG, et al. Temporal association of changes in fasting blood glucose and body mass index with diagnosis of pancreatic cancer. *Am J Gastroenterol* 2009; 104: 2318–2325.
90. Bonovas S, Filioussi K, Tsantes A. Diabetes mellitus and risk of prostate cancer: a meta-analysis. *Diabetologia* 2004; 47: 1071–1078.
91. Kasper JS, Giovannucci E. A meta-analysis of diabetes mellitus and the risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2006; 15: 2056–2062.
92. Shaneyfelt T, Husein R, Bubley G, Mantzoros CS. Hormonal predictors of prostate cancer: a meta-analysis. *J Clin Oncol* 2000; 18: 847–853.
93. Kalyani RR, Dobs AS. Androgen deficiency, diabetes, and the metabolic syndrome in men. *Curr Opin Endocrinol Diabetes Obes* 2007; 14: 226–234.
94. Reaven GM. Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes* 1988; 37: 1595–1607.
95. Grossmann M, Thomas MC, Panagiotopoulos S, et al. Low testosterone levels are common and associated with insulin resistance in men with diabetes. *J Clin Endocrinol Metab* 2008; 93: 1834–1840.
96. Stocks T, Lukanova A, Rinaldi S, et al. Insulin resistance is inversely related to prostate cancer: a prospective study in Northern Sweden. *Int J Cancer* 2007; 120: 2678–2686.
97. Baradaran N, Ahmadi H, Salem S, et al. The protective effect of diabetes mellitus against prostate cancer: role of sex hormones. *Prostate* 2009; 69: 1744–1750.
98. Ma J, Li H, Giovannucci E, et al. Prediagnostic body-mass index, plasma C-peptide concentration, and

prostate cancer-specific mortality in men with prostate cancer: a long-term survival analysis. *Lancet Oncol* 2008; 9: 1039–1047.

99. Aguilar-Bryan L, Nichols CG, Wechsler SW, et al. Cloning of the beta cell high-affinity sulfonylurea receptor: a regulator of insulin secretion. *Science* 1995; 268: 423–426.
100. Malhi H, Irani AN, Rajvanshi P, et al. KATP channels regulate mitogenically induced proliferation in primary rat hepatocytes and human liver cell lines. Implications for liver growth control and potential therapeutic targeting. *J Biol Chem* 2000; 275: 26050–26057.
101. Wondergem R, Cregan M, Strickler L, et al. Membrane potassium channels and human bladder tumor cells: II. Growth properties. *J Membr Biol* 1998; 161: 257–262.
102. Qian X, Li J, Ding J, et al. Glibenclamide exerts an antitumor activity through reactive oxygen species-c-jun NH2-terminal kinase pathway in human gastric cancer cell line MGC-803. *Biochem Pharmacol* 2008; 76: 1705–1715.
103. Houghton PJ, Houghton JA. Antitumor diarylsulfonylureas: novel agents with unfulfilled promise. *Invest New Drugs* 1996; 14: 271–280.
104. Lee CW, Hong DH, Han SB, et al. A novel stereo-selective sulfonylurea, 1-[1-(4-aminobenzoyl)-2,3-dihydro-1H-indol-6-sulfonyl]–4-phenyl-imidazolidin-2-one, has antitumor efficacy in in vitro and in vivo tumor models. *Biochem Pharmacol* 2002; 64: 473–480.
105. Bowker SL, Majumdar SR, Veugelers P, Johnson JA. Increased cancer-related mortality for patients with type 2 diabetes who use sulfonylureas or insulin. *Diabetes Care* 2006; 29: 254–258.
106. Donadon V, Balbi M, Ghersetti M, et al. Antidiabetic therapy and increased risk of hepatocellular carcinoma in chronic liver disease. *World J Gastroenterol* 2009; 15: 2506–2511.
107. Currie CJ, Poole CD, Gale EA. The influence of glucose-lowering therapies on cancer risk in type 2 diabetes. *Diabetologia* 2009; 52: 1766–1777.
108. Kurtzhals P, Schaffer L, Sorensen A, et al. Correlations of receptor binding and metabolic and mitogenic potencies of insulin analogs designed for clinical use. *Diabetes* 2000; 49: 999–1005.
109. Mayer D, Shukla A, Enzmann H. Proliferative effects of insulin analogues on mammary epithelial cells. *Arch Physiol Biochem* 2008; 114: 38–44.
110. Weinstein D, Simon M, Yehezkel E, et al. Insulin analogues display IGF-I-like mitogenic and anti-apoptotic activities in cultured cancer cells. *Diabetes Metab Res Rev* 2009; 25: 41–49.
111. Stammberger I, Bube A, Durchfeld-Meyer B, et al. Evaluation of the carcinogenic potential of insulin glargine (LANTUS) in rats and mice. *Int J Toxicol* 2002; 21: 171–179.
112. Hemkens LG, Grouven U, Bender R, et al. Risk of malignancies in patients with diabetes treated with human insulin or insulin analogues: a cohort study. *Diabetologia* 2009; 52: 1732–1744.
113. Colhoun HM. Use of insulin glargine and cancer incidence in Scotland: a study from the Scottish Diabetes Research Network Epidemiology Group. *Diabetologia* 2009; 52: 1755–1765.
114. Giardiello FM, Brensinger JD, Tersmette AC, et al. Very high risk of cancer in familial Peutz-Jeghers syndrome. *Gastroenterology* 2000; 119: 1447–1453.
115. Zhou G, Myers R, Li Y, et al. Role of AMP-activated protein kinase in mechanism of metformin action. *J Clin Invest* 2001; 108: 1167–1174.
116. Hawley SA, Boudeau J, Reid JL, et al. Complexes between the LKB1 tumor suppressor, STRAD alpha/beta and MO25 alpha/beta are upstream kinases in the AMP-activated protein kinase cascade. *J Biol* 2003; 2: 28.
117. Cazzaniga M, Bonanni B, Guerrieri-Gonzaga A, Decensi A. Is it time to test metformin in breast cancer clinical trials? *Cancer Epidemiol Biomarkers Prev* 2009; 18: 701–705.
118. Zakikhani M, Dowling R, Fantus IG, et al. Metformin is an AMP kinase-dependent growth inhibitor for breast cancer cells. *Cancer Res* 2006; 66: 10269–10273.
119. Zakikhani M, Dowling RJ, Sonenberg N, Pollak MN. The effects of adiponectin and metformin on prostate and colon neoplasia involve activation of AMP-activated protein kinase. *Cancer Prev Res* 2008; 1: 369–375.
120. Hirsch HA, Iliopoulos D, Tsichlis PN, Struhl K. Metformin selectively targets cancer stem cells, and acts together with chemotherapy to block tumor growth and prolong remission. *Cancer Res* 2009; 69: 7507–7511.
121. Ben Sahra I, Laurent K, Loubat A, et al. The antidiabetic drug metformin exerts an antitumoral effect in vitro and in vivo through a decrease of cyclin D1 level. *Oncogene* 2008; 27: 3576–3586.
122. Anisimov VN, Berstein LM, Egormin PA, et al. Effect of metformin on life span and on the development of spontaneous mammary tumors in HER-2/neu transgenic mice. *Exp Gerontol* 2005; 40: 685–693.
123. Buzzai M, Jones RG, Amaravadi RK, et al. Systemic treatment with the antidiabetic drug metformin selectively impairs p53-deficient tumor cell growth. *Cancer Res* 2007; 67: 6745–6752.
124. Algire C, Zakikhani M, Blouin MJ, et al. Metformin attenuates the stimulatory effect of a high-energy diet on in vivo LLC1 carcinoma growth. *Endocr Relat Cancer* 2008; 15: 833–839.
125. Evans JM, Donnelly LA, Emslie-Smith AM, et al. Metformin and reduced risk of cancer in diabetic patients. *BMJ* 2005; 330: 1304–1305.
126. Jiralerspong S, Palla SL, Giordano SH, et al. Metformin and pathologic complete responses to neoadjuvant chemotherapy in diabetic patients with breast cancer. *J Clin Oncol* 2009; 27: 3297–3302.
127. Saez E, Rosenfeld J, Livolsi A, et al. PPAR gamma signaling exacerbates mammary gland tumor development. *Genes Dev* 2004; 18: 528–540.
128. Tontonoz P, Singer S, Forman BM, et al. Terminal differentiation of human liposarcoma cells induced by ligands for peroxisome proliferator-activated receptor gamma and the retinoid X receptor. *Proc Natl Acad Sci U S A* 1997; 94: 237–241.
129. DuBois RN, Gupta R, Brockman J, et al. The nuclear eicosanoid receptor, PPARgamma, is aberrantly expressed in colonic cancers. *Carcinogenesis* 1998; 19: 49–53.
130. Mueller E, Smith M, Sarraf P, et al. Effects of ligand activation of peroxisome proliferator-activated

receptor gamma in human prostate cancer. *Proc Natl Acad Sci U S A* 2000; 97: 10990–10995.

131. Sarraf P, Mueller E, Jones D, et al. Differentiation and reversal of malignant changes in colon cancer through PPARgamma. *Nat Med* 1998; 4: 1046–1052.
132. Mueller E, Sarraf P, Tontonoz P, et al. Terminal differentiation of human breast cancer through PPAR gamma. *Mol Cell* 1998; 1: 465–470.
133. Elstner E, Muller C, Koshizuka K, et al. Ligands for peroxisome proliferator-activated receptor gamma and retinoic acid receptor inhibit growth and induce apoptosis of human breast cancer cells in vitro and in BNX mice. *Proc Natl Acad Sci U S A* 1998; 95: 8806–8811.
134. Lefebvre AM, Chen I, Desreumaux P, et al. Activation of the peroxisome proliferator-activated receptor gamma promotes the development of colon tumors in C57BL/6J-APCMin/+mice. *Nat Med* 1998; 4: 1053–1057.
135. Saez E, Tontonoz P, Nelson MC, et al. Activators of the nuclear receptor PPARgamma enhance colon polyp formation. *Nat Med* 1998; 4: 1058–1061.
136. Dormandy JA, Charbonnel B, Eckland DJ, et al. Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (Prospective Pioglitazone Clinical Trial In Macrovascular Events): a randomised controlled trial. *Lancet* 2005; 366: 1279–1289.
137. Govindarajan R, Ratnasinghe L, Simmons DL, et al. Thiazolidinediones and the risk of lung, prostate, and colon cancer in patients with diabetes. *J Clin Oncol* 2007; 25: 1476–1481.
138. Burstein HJ, Demetri GD, Mueller E, et al. Use of the peroxisome proliferator-activated receptor (PPAR) gamma ligand troglitazone as treatment for refractory breast cancer: a phase II study. *Breast Cancer Res Treat* 2003; 79: 391–397.
139. Blanquicett C, Roman J, Hart CM. Thiazolidinediones as anti-cancer agents. *Cancer Ther* 2008; 6: 25–34.

SPECIAL FEATURE

Colonoscopy Versus Computed Tomography Colonography for Colorectal Cancer Screening

Reza Akhtar, MD, Michelle Lee, MD, and Steven H. Itzkowitz, MD

Mount Sinai School of Medicine, New York, NY

OUTLINE

- Optical Colonoscopy
    - **Protection Offered by Optical Colonoscopy**
    - **Accuracy**
    - **Reasons for Misses: Adenoma Type, Colonoscopy Quality, and Technology**
    - **Complications**
    - **Interval**
- Computed Tomography Colonography
    - **Accuracy**
    - **Reasons for Misses: Quality and Technology**
    - **Complications**
- Summary

ABSTRACT

Colorectal cancer is second only to lung cancer as the leading cause of death among North Americans of both sexes. Although screening rates for colorectal cancer in the United States have increased over the past decade, these rates (in the range of 45%–60%) are still lower than the screening rates for breast cancer (approximately 80%). Optical colonoscopy has been recognized as the preferred method for colorectal cancer screening in the United States, but computed tomography colonography has recently been gaining favor. This article compares the 2 methods with respect to both advantages and disadvantages. ***Mt Sinai J Med* 77:214–224, 2010.** © *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Steven H. Itzkowitz
Mount Sinai School of Medicine
New York, NY
Email: steven.itzkowitz@mountsinai.org.

Key Words: cancer screening, colonoscopy, colorectal cancer, computed tomography colonography.

Colorectal cancer (CRC) afflicts more than 140,000 men and women in the United States each year.[1] It is second only to lung cancer as the leading cause of death from cancer among both sexes in North America. Fortunately, despite these deadly statistics, the incidence of CRC in both men and women has declined since the 1990s.[2] Increased physician and patient awareness, better compliance with screening guidelines, and thus earlier and more effective treatment of precancerous colorectal lesions have been implicated. Although screening rates for CRC in the United States have increased over the past decade,[3,4] these rates (in the range of 45%–60%) are still lower than the screening rates for breast cancer (approximately 80%).[5] This suggests that there is still room to improve CRC screening rates. In the most

Although screening rates for CRC in the United States have increased over the past decade, these rates (in the range of 45%–60%) are still lower than the screening rates for breast cancer (approximately 80%). This suggests that there is still room to improve CRC screening rates.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20175

© 2010 Mount Sinai School of Medicine

recent set of guidelines, the US Multisociety Task Force endorsed the use of stool-based tests (fecal occult blood tests and stool DNA tests) as well as sigmoidoscopy, optical colonoscopy (OC), barium enema, and computed tomography colonography (CTC) as viable options for CRC screening.[6] In contrast, the latest guidelines of the US Preventive Services Task Force report that there is insufficient evidence to recommend CTC or stool DNA testing at this time.[7] The indeterminate status of CTC was based largely on the task force's inability to assess the potential harm of radiation and extracolonic findings. In practice, OC has become increasingly recognized as the preferred method for CRC screening in the United States, and it has earned the endorsement of several organizations such as the American College of Gastroenterology[8] and the New York City Department of Health and Mental Hygiene.[9] Indeed, OC has the singular advantage of being both diagnostic and therapeutic while maintaining high accuracy (Table 1). Despite these advantages, OC has faced scrutiny for missing lesions and for other barriers such as the need for sedation, time off work, and bowel cleansing and procedural complications. Thus, efforts to improve OC have paralleled the search for an alternative screening test. CTC has recently gained favor as a structural examination of the colorectum. First introduced by Vining *et al.*[10] at the 1994 annual meeting of the Society of Gastrointestinal Radiologists, CTC has since been the focus of many investigations, and many possible advantages over OC have been suggested. First, CTC is relatively noninvasive and does not require sedation (Table 1). Second, it can be performed rather quickly and appears to have a lower risk of procedure-related complications.[11,12] Other desirable factors are the potential for cost savings and perhaps greater acceptability by the general population. This latter point is of considerable importance because of the unacceptably low compliance rates with current screening recommendations.[13] Limitations of CTC include the need for bowel preparation, some degree of patient discomfort from air insufflation of the colon, the relative lack of sensitivity for polyps smaller than 5 mm, the need to undergo OC if a lesion is found, the exposure to radiation, and the debate about the significance and cost/benefit of finding extracolonic pathology. The purpose of this article is to compare and contrast OC and CTC for CRC screening. Table 1 summarizes the main features of each procedure.

OC has become increasingly recognized as the preferred method for CRC screening in the United States because OC has the singular advantage of being both diagnostic and therapeutic while maintaining high accuracy. Despite these advantages, OC has faced scrutiny for missing lesions and for other barriers such as the need for sedation, time off work, and bowel cleansing and procedural complications.

CTC has recently gained favor as a structural examination of the colorectum because it is relatively noninvasive and does not require sedation. It can be performed rather quickly and appears to have a lower risk of procedure-related complications. Other desirable factors are the potential for cost savings and perhaps greater acceptability by the general population.

OPTICAL COLONOSCOPY

Curing a malignant lesion is the obvious goal of most therapeutic oncological endeavors. More enticing is the ability to confidently identify and treat a premalignant or early malignant lesion and thereby further reduce the chance of spread and subsequent morbidity and mortality. Most CRCs arise from adenomatous polyps.[14] As such, the adenomatous polyp represents an important target for CRC screening.

Protection Offered by Optical Colonoscopy

Many studies have demonstrated the protective effects of OC and have justified the colonoscopic removal of adenomatous polyps as a widely accepted form of cancer prevention and treatment. However,

Table 1. *Comparison of Optical Colonoscopy and Computed Tomography Colonography.*

	Optical Colonoscopy	Computed Tomography Colonography
Utility	Diagnostic and therapeutic	Therapeutic only
Accuracy	Highly accurate	Highly accurate
Interval	Every 10 years	Uncertain
Bowel preparation needed?	Yes	Yes; fecal tagging improves results.
Sedation needed?	Yes (in most centers)	No
Complications	Perforation and bleeding (rare)	Perforation (very rare)
Miss rate	Up to 27% of polyps	Up to 50% of polyps < 5 mm
Reimbursement	Covered by most insurance plans	Limited coverage (noncoverage decision by the Centers for Medicare and Medicaid Services, May 2009)
Other disadvantages	Absence from work Poor patient acceptance in some settings	Radiation exposure Detection of clinically unimportant extracolonic lesions may result in added expense and morbidity. Need for optical colonoscopy if it is positive Patient acceptance not yet known
Other advantages	May detect lesions missed by computed tomography colonography Cost-effective	May detect clinically important extracolonic pathology May detect lesions missed by optical colonoscopy Cost-effective in some analyses

Many studies have demonstrated the protective effects of OC and have justified the colonoscopic removal of adenomatous polyps as a widely accepted form of cancer prevention and treatment.

the extent to which this method prevents CRC varies in the literature in both degree and length of protection.

The landmark National Polyp Study (NPS) demonstrated that removing adenomatous precursor lesions reduced the expected incidence of CRC by an impressive degree: a 76% to 90% reduction over a mean follow-up of 5.9 years in comparison with 3 reference groups.[15] Likewise, outside the setting of a clinical trial, a 66% reduction in CRC incidence in the 10 years after adenoma removal was reported by an Italian group on the basis of routine clinical practice.[16] In addition, long-term follow-up of the NPS confirmed a sustained impact of polypectomy on incidence and mortality.[17] In contrast, other studies of adenoma cohorts have not found the dramatic reduction in incident cancer rates seen in the NPS.[18] The reasons for this are unclear but include possible superior baseline clearing colonoscopy and exclusion of adenomas larger than 3 cm in the NPS and different study designs. Mathematical modeling has suggested that most of the reduction in CRC incidence and mortality in the NPS resulted from the initial, baseline colonoscopy, and this emphasizes the need for high-quality examinations (as discussed later).[19] Despite these somewhat conflicting reports, more recent evidence for the protective effect of colonoscopy comes from a follow-up study of 715 patients who underwent screening colonoscopy between 1989 and 1993; the observed CRC incidence and mortality rates were compared to Surveillance, Epidemiology, and End Results data.[20] In this group of average-risk patients, only 26% of whom had an adenoma at the initial screening colonoscopy, there was a 67% relative risk reduction of CRC incidence as well as a 65% reduction in death from CRC. Furthermore, there is evidence from an endoscopic claims database study in Manitoba, Canada that the risk of developing CRC remains decreased for more than 10 years after a negative colonoscopy.[21]

Recent attention has been focused on whether OC prevents cancer in all segments of the colon. For example, the Manitoba study found that the proportion of CRC located in the proximal colon was higher in the cohort that had a negative colonoscopy than in the general population of Manitoba.[21] Another Canadian investigation found that although the overall CRC risk was decreased after OC [odds ratio = 0.69, 95% confidence interval (CI) = 0.63–0.74], there was no decreased risk of right-sided cancers (odds ratio = 1.07, 95% CI = 0.94–1.21).[22] The latter study has been criticized because the data came from a claims database that does not permit an analysis of whether all colonoscopies reached the cecum, whether the indication was screening versus a diagnostic examination, and whether there were possible unmeasured confounders; moreover, only 31% of the colonoscopies were performed by gastroenterologists. Nonetheless, this study has stirred debate about the accuracy of OC, the need to account for the quality of the examination, and the possibility that right-sided neoplasia may have a different biology than distal disease.

Accuracy

Although OC is commonly used as the gold-standard CRC screening test, variability in CRC prevention with screening OC can be explained, in part, by the fact that even in skilled hands, OC misses lesions.

Although OC is commonly used as the gold-standard CRC screening test, variability in CRC prevention with screening OC can be explained, in part, by the fact that even in skilled hands, OC misses lesions.

Quantifying the miss rate for screening OC can be challenging. Miss rates have been estimated by re-examination of the colon and rectum either by tandem colonoscopy on the same day or by short-interval follow-up colonoscopy. The latter strategy assumes that new and particularly larger lesions on surveillance were missed on the first pass instead of newly arising in the interim. In addition, CTC studies have highlighted lesions that were missed by OC.

By comparing adenoma detection rates between same-day consecutive colonoscopies, Rex and colleagues[23] found miss rates as high as 27% for adenomas ≤ 5 mm, 13% for 6- to 9-mm adenomas, and 6% for adenomas ≥ 1 cm. Imperiale and colleagues,[24] however, found a low incidence of advanced adenomas at an average of 5 years after a negative baseline OC and calculated that approximately 79 average-risk patients would need to be surveyed at 5 years to find 1 advanced adenoma. Rex and colleagues[25] performed an earlier study with similar findings. Five years after baseline screening OC, the overall incidence of adenomas was 27%, and most (90%) were less than 5 mm; only 1 polyp was greater than 1 cm, and no cancer or advanced neoplasia was found. In a review of 6 studies investigating 465 patients, van Rijn and colleagues[26] described the following miss rates by tandem OC: 2.1% for adenomas ≥ 10 mm (95% CI = 0.3%–7.3%), 13% for 5- to 10-mm adenomas (95% CI = 8.0%–18%), and 26% for 1- to 5-mm adenomas (95% CI = 27%–35%).

CTC, being a much better structural examination than a barium enema, has also enlightened us about missed lesions during OC. Among 1233 asymptomatic adults who underwent same-day OC and CTC, Pickhardt and colleagues[27] found that of 511 polyps detected by OC, 55 (10.8%) were found only on second-look OC after segmental unblinding of CTC results. Adenomas missed by OC were usually on the proximal side of a fold or, ironically, at the anal verge.

Reasons for Misses: Adenoma Type, Colonoscopy Quality, and Technology

Overall, we are left with some conflicting evidence regarding the protective effects of screening OC. This has directed attention toward improving the quality of screening OC across the country. The accuracy of polyp detection in clinical studies may not adequately reflect that of the community. Because of the varied experience of colonoscopists and because some specialize particularly in endoscopy, these standards have become a focal point of improvements to prevent suboptimal examination of the colon and rectum. Additionally, it is important to consider the clinical outcome of missed lesions to determine their significance in a patient's overall health status.

Malignant potential differs according to the characteristics of a polyp, such as its size. Advanced adenomas are a more desirable target for intervention because the progression to carcinoma is more likely than with small adenomas.[28] Advanced adenomas are typically defined as tubular adenomas with more than 25% villous component, adenomas greater than or equal to 1 cm in size, high-grade dysplasia, or early invasive cancer. Of course, a larger polyp will likely be more easily detected by a structural investigation, although this is never guaranteed. The exact incidence of cancer in patients with advanced adenomas is not known.[29] More recent data have drawn attention to other features of polyps that may also factor into their malignant potential. These include serrated adenomas, which are believed to progress to malignancy in a different fashion than traditional adenomas,[30] and flat (or nonpolypoid) adenomas. The risk of carcinoma in flat adenomas ranges from 0% to 7% in published studies.[31,32] A study by Soetikno and colleagues[33] of 1819 consecutive veterans in the Palo Alto Veterans

Affairs system undergoing screening, surveillance, or diagnostic OC reported that the overall prevalence of nonpolypoid colorectal neoplasm was relatively high at 9.3%; the odds ratio of a neoplasm was 3.25 (95% CI = 2.19 – 4.85) for nonpolypoid adenomas versus polypoid lesions. Others have found higher neoplastic potential for flat adenomas.[34] However, when O'Brien and colleagues[35] reclassified sessile adenomas as flat adenomas from the NPS, they found no increased risk for high-grade dysplasia or advanced histology at either baseline or surveillance OC. Thus, although the precise natural history of flat adenomas continues to be elucidated, these lesions are important to consider when the accuracy of screening tests in detecting polyps is being evaluated.

Much attention has also been given to the quality of screening OC. These measures depend on both patient and operator characteristics. Important variables include the adequacy of bowel preparation, the ability to reach the cecum, sufficient withdrawal time, and operator experience. Nonadherence to these established guidelines greatly influences test efficacy. Poor bowel preparation leads to increased costs and repeat procedures at intervals earlier than those recommended by current guidelines.[36] Patients often cite the bowel preparation as the most unpleasant part of the entire procedure. Sedation aids in comfort during OC, although these medications can add to the overall risk of the procedure and require patients to take time off work and use chaperones to ensure a safe return home. Careful mucosal examination also requires adequate time of inspection. Since 2002, guidelines have recommended a 6- to 10-minute withdrawal phase to enhance polyp detection.[6] In their own community practice, Barclay and colleagues[37] found that practitioners employing withdrawal times of at least 6 minutes had higher detection rates of any neoplasia and advanced neoplasia. Simmons and colleagues[38] examined over 10,000 procedures performed by 43 endoscopists in an academic outpatient clinic and found that longer withdrawal times strongly correlated with higher polyp detection rates even beyond a median of 6.7 minutes.

Optical imaging technology has evolved significantly over the last decade, and the details, magnification, and extent of what is seen during OC have been improved. Although it is tempting to suppose that better imaging will lead to higher adenoma detection rates, this has not been borne out in the literature.[39–41] Narrow band imaging (NBI) allows for detection of subtle tissue structures through optical filtering using green and blue light bands instead of white ones. A pilot feasibility study found that, in 40 patients, NBI visualized 51 more polyps, more than half of which were adenomatous on histological confirmation.[42] However, in a larger, randomized population, no difference was found in adenoma detection between white-light endoscopy and NBI.[39] An early study on autofluorescence imaging found a 20% reduction in the adenoma miss rate in the right colon versus traditional white-light endoscopy.[43] However, this was a single-practitioner study, and the absolute miss rate in the white-light endoscopy group was high at 49%. Mucosal enhancement by total colonic dye spray may increase the detection of diminutive adenomas[44]; however, the procedure takes twice as long as the traditional technique, and the clinical significance of finding such diminutive lesions is uncertain.

Complications

Prior to the start of the procedure, informed consent is obtained from the patient, who is told the risks and benefits of OC and its alternatives. Of the risks mentioned, bleeding and perforation tend to invoke the most alarm in a patient's eyes. Similar to previously published US population–based studies, a recent Canadian study of screening OC in an average-risk population reported absolute risks of bleeding and perforation of 1.64/1000 and 0.85/1000, respectively.[45] The risk of death is unfortunately not zero, although it is very low at 0.074/1000, or 1 in 14,000. These risks are higher with older age, polypectomy during the procedure, and lesser experienced endoscopists. OC carries a higher risk of perforation than sigmoidoscopy.[46] This is an important consideration for physicians when they are deciding which screening measure to recommend for a patient who is at higher risk for OC-related complications.

Interval

In the United States, national guidelines for screening have remained largely unchanged. Initially, the intervals for screening were based on studies conducted over 15 years ago that used sigmoidoscopy for the reduction of mortality for distal CRC at 10 years.[47,48] Despite the limitations of these studies (ie, the study design and the lack of full colonic evaluation), the published timeframe of 10 years remains the recommended screening interval for surveillance examinations in patients with a normal baseline colonoscopic examination.[6,7]

The timing of surveillance OC is dependent on the findings of the initial examination. Almost 2 decades ago, the NPS prospectively compared the surveillance of patients with 1 or more adenomas at 1 and 3 years versus surveillance at 3 years only and found no difference in the rate of recurrent adenomas.[19] As in this study, subsequent analyses have lent support to longer screening intervals on the basis of the incidence rates of adenomatous polyps after negative index colonoscopies. These rates vary widely from 16% to 41% and mostly predate the 21st century.[25,50,51]

The characteristics of both the patient and polyp have been shown to influence the risk for future recurrence of polyps and CRC. In a recent pooled analysis of 8 large, prospective studies, Martinez and colleagues[52] described the risk of a metachronous advanced adenoma or CRC after initial OC with polypectomy as a function of the baseline patient and polyp characteristics. Noting that the strongest predictors of risk were patient age and prior polyp size, the authors concluded that 1 in 10 patients will develop an advanced colorectal neoplasm on surveillance OC, the minority of which will be cancer (1 in 150). Interestingly, the proximal location of polyps was also found to be a risk factor, and this speaks to the possible disparity in tumor biology between left-sided and right-sided CRCs.

COMPUTED TOMOGRAPHY COLONOGRAPHY

Accuracy

During the establishment of CTC as a screening technology, earlier studies were rather heterogeneous in design. They often included symptomatic or higher risk patients in order to optimize lesion detection, less sophisticated computed tomography (CT) scanners and computer software were available, and radiologists were still learning to read these new imaging studies. In 2005, Mulhall and colleagues[53] synthesized the literature up to that time in a meta-analysis of prospective studies comparing polyp detection by CTC with standard OC or surgery as the gold standard. The per-patient sensitivity for CTC ranged from 21% to 91%, with an overall pooled sensitivity of 70% (95% CI = 53% – 87%). The sensitivity increased directly in proportion to polyp size: 48% for polyps $\leq$ 6 mm (95% CI = 25% – 70%), 70% for 6- to 9-mm polyps (95% CI = 55% – 84%), and 85% for polyps $\geq$ 9 mm (95% CI = 79% – 91%). Reasons for the observed heterogeneity were related to technology: studies using thinner slices for imaging appeared to have better sensitivity. Additionally, studies using 3-dimensional imaging and fly-through techniques had consistently higher sensitivities, which approached 95%. In contrast to per-patient sensitivity, the specificity was more homogeneous over various polyp sizes; the overall specificity was 86%. The authors hypothesized that with the refinement of CT scanners, improved patient bowel preparation, and evolving software for CTC, the diagnostic accuracy would likely improve.

The landmark study that demonstrated CTC to be a potentially useful CRC screening test came in 2003 when Pickhardt and colleagues[12] studied an asymptomatic screening population. These investigators examined the accuracy of CTC and used an enhanced gold standard to describe their results. During OC, after the examination of each colonic segment, the practitioner was provided the results of the patient's previously performed CTC for that particular segment. This segmental unblinding allowed for a better assessment of false negatives by OC. The sensitivity of CTC for detecting lesions was 92.2% for adenomatous polyps $\geq$ 10 mm (95% CI = 81.1–97.8), 92.6% for polyps $\geq$ 8 mm (95% CI = 85.4–97.0), and 85.7% for adenomatous polyps $\geq$ 6 mm (95% CI = 80.2–90.1), in contrast to rates of 88.2% (95% CI = 76.1–95.6), 89.5% (95% CI = 81.5–94.8), and 90.0% (95% CI = 85.1–93.7), respectively, for initial OC. Additionally, 55 polyps $\geq$ 5 mm in diameter that were missed by OC were picked up by CTC; 1 polyp was 11 mm and malignant. The authors concluded that CTC not only has high sensitivity but also maintains acceptable specificity for adenomas > 6 mm in diameter. Using a different study design, Kim and colleagues[54] more recently compared detection rates for advanced neoplasia in a group of patients who underwent primary CTC screening versus primary OC at a single institution. In this nonrandomized study, the total number of advanced neoplasms was comparable between the 2 groups. Advanced neoplasia was confirmed in 3.2% of the CTC group and in 3.4% of the OC group. Although the 2 screening modalities had similar detection rates for advanced neoplasms, CTC resulted in fewer complications and fewer polypectomies as smaller lesions were not referred for colonoscopic removal. The overall prevalence of advanced neoplasia in the entire group was 3.3%, which is lower than the prevalence of 4% to 6% seen

> *CTC not only has high sensitivity but also maintains acceptable specificity for adenomas > 6 mm in diameter.*

in some other studies.[12,37,55–57] The authors attributed this difference to variations in age, sex, ethnic and family background, and frequency of symptoms. It is worth noting that 7.9% of the CTC group members were referred for OC because of CTC findings, and this provides some insight into the potential impact on colonoscopy volume if CTC use were to become more widespread.

To build on the findings from research centers with a particular interest in CTC, Johnson and colleagues[58] conducted an important study in asymptomatic individuals to define the accuracy of CTC as a CRC screening tool in more of a real-world setting. In this large, prospective, multicenter study known as the American College of Radiology Imaging Network 6664 (ACRIN-6664) study, each participant underwent CTC, which was followed by blinded OC. When lesions ≥ 10 mm were detected on CTC but not on OC, patients were instructed to have a repeat OC within 90 days. The endoscopist was made aware of the CTC results before the repeat OC. The OC results (including the repeat OC results) and the histological examination of the biopsy samples were the reference standard for determining the lesion size, location, and histological type. The mean (+/− standard deviation) per-patient estimate for the sensitivity, specificity, positive predictive value, and negative predictive value of detecting large adenomas and cancer by CTC were 0.90 (+/−0.03), 0.86 (+/−0.02), 0.23 (+/−0.02), and 0.99 (+/−0.01), respectively. The sensitivity of CTC for the detection of adenomas or cancers ≥ 5, 6, 7, 8, and 9 mm was 0.65 (95% CI = 0.58–0.73), 0.78 (95% CI = 0.71–0.85), 0.84 (95% CI = 0.78–0.91), 0.87 (95% CI = 0.80–0.93), and 0.90 (95% CI = 0.83–096), respectively, with the specificity ranging from 0.86 to 0.89.

Reasons for Misses: Quality and Technology

As one might expect, the accuracy of CTC, like that of OC, relies on a sufficiently cleansed and adequately distended colon. Other technological factors are important, such as rapid thin-section CT examination in both supine and prone positions, computerized processing of the data to create 2-dimensional and 3-dimensional CTC images, and experienced readers.[59] An additional consideration for CTC is residual fluid in the lumen. Residual fluid can mask polyps by decreasing the contrast between the bowel wall and the colonic lumen. To further increase the diagnostic yield, an iodinated oral contrast agent is administered prior to the scan in a procedure known as fluid tagging. Fluid tagging is believed to increase the attenuation of residual fluid and allow for better visualization of polyps that are submerged in intraluminal fluid.[59] Additionally, patients may be instructed to ingest barium over several days; this tags the stool (fecal tagging) and subsequently allows further contrast of solid luminal contents against the colonic mucosa.

The accuracy of CTC, like that of OC, relies on a sufficiently cleansed and adequately distended colon. Other technological factors are important, such as rapid thin-section CT examination in both supine and prone positions, computerized processing of the data to create 2-dimensional and 3-dimensional CTC images, and experienced readers.

After these measures, thin-section abdominopelvic CT is performed with the patient in both supine and prone positions. Generally, a multidetector CT scanner is preferred over single-detector scanners because of its ability to create thinner slices. This increases polyp detection sensitivity by 5%.[53] After the images are acquired, computer software creates both 2-dimensional and 3-dimensional images, which ideally are evaluated by an experienced radiologist.

According to the American College of Radiology, practice-based quality metrics for CTC have been in place since 2007. These include process measures of the CTC technique, image quality, patient preparation and outcome measures such as rates of true positives, colonic perforation, and extracolonic findings.[6] The ACRIN-6664 study reported that half of the radiologists did not pass the initial certifying examination (after either 1.5 days of training or experience of 500 or more cases), although all readers did pass after additional training.[58] Additionally, the American College of Radiology also reported the construction of an interactive training facility for CTC that provides for the evaluation of individual certification and proficiency. Other issues to consider with respect to the quality of CTC are CT measurement approaches, viewing displays, and review times for CTC similar to withdrawal time benchmarks for OC.

Complications

There remains much uncertainty about the potential harm associated with CTC. One cause for concern is the related radiation exposure. One study estimated that the absolute lifetime cancer risk associated with the radiation exposure from typical current scanner techniques is approximately 0.14% for a 50-year-old and about half that for a 70-year-old.[60] A CTC examination in a 50-year-old was estimated to add 0.044% more to the lifetime risk of colon cancer. Another potential procedure-related harm of CTC is perforation due to the need to insufflate the colon with air to achieve adequate visualization. However, the overall risk of perforation appears to be low, with reported rates of symptomatic perforation ranging from 0.03% (1 in 3400 patients)[2] to 0.06% (1 in 7000 patients).[61]

The implications of extracolonic findings are unclear because having knowledge of silent abnormalities outside the colon can be both advantageous and disadvantageous.

A unique situation with CTC, in comparison with OC, is that CTC can detect abnormalities outside the colon. The implications of extracolonic findings are unclear because having knowledge of silent abnormalities outside the colon can be both advantageous and disadvantageous. In an asymptomatic screening population, the incidence of unsuspected but potentially crucial extracolonic findings is approximately 4.5%.[6] In the study by Johnson and colleagues,[58] extracolonic findings were detected in 66% of the participants, but only 16% required additional investigation or urgent care. Although there is no doubt that there are potential benefits from these discoveries, there are associated risks and costs for both the patient and the healthcare system that will need to be considered, especially when the findings are false positives. Evaluating a false positive by other tests or surgery can result in morbidity and potential mortality, which would detract from the overall merits of a screening program. In recent years, some gastroenterologists have taken training courses in how to read CTC studies. Although the interpretation of intraluminal findings may come easily, issues are raised about the reliability of nonradiologists in interpreting extracolonic findings (prompting the development of software programs that permit visualization only of the intraluminal gastrointestinal tract) as well as reimbursement.

A major factor that could make CTC a favorable screening tool is its potential for widespread acceptability by the general population. However, in one study, CTC was associated with greater discomfort than OC.[12] Although 68.3% of patients rated CTC as more acceptable than OC (because of overall convenience), only 50% of the patient stated that they would choose CTC as their choice for future screening tests versus 41% who chose OC. The authors stated there was an unintended bias favoring OC, in that some of the patients who had polyp removal assumed that the polypectomies were necessary. At present, reimbursement for CTC is covered in the setting of incomplete colonoscopy, in the setting of contraindications to colonoscopy, and in some high-risk settings. In research settings, OC is typically performed on the same day as CTC to avoid a second preparation. However, in a real-world setting, the logistics of coordinating these 2 examinations on the same day (or even the next day) are formidable.

There is debate about whether all polyps, especially ones smaller than 5 mm, need to be removed or even reported.

Finally, there is debate about whether all polyps, especially ones smaller than 5 mm, need to be removed or even reported. Many radiologists have suggested that diminutive polyps (smaller than 5 mm) not be reported because these lesions typically fall below the threshold for accurate detection by CTC. Indeed, only about 1.7% of colonoscopically removed polyps of this size will have advanced histology, and cancers are quite rare.[62] More controversial is how to manage polyps ranging in size from 6 to 9 mm. These lesions are referred for either colonoscopy or repeat CTC in 2 to 3 years. However, if the decision is made to wait 2 to 3 years, patients should be informed that up to 6.6% of lesions of this size may already demonstrate advanced histology.[62] Pickhardt and colleagues[63] constructed a decision analysis model to determine the cost effectiveness of using CTC surveillance for small polyps found on CTC versus immediate polypectomy. Their study found that using 3-year CTC surveillance of 6- to 9-mm polyps was more cost-effective, with immediate polypectomy costing in excess of $372,853 per life year gained. Additionally, they estimated that the 5-year CRC death rate from an unresected small polyp (0.007%) was comparable to the rate of death resulting from OC-related complications (0.006%). The current recommendation for referring a patient for OC is a polyp found on CTC that is 6 mm or greater in size.[6] This threshold is a result not of clinical studies but rather of expert recommendation. This would make OC referral relatively common because as many as 1 in 3 persons and as few as 1 in 8 persons would require colonoscopy.[13]

SUMMARY

Direct endoscopic visualization of the colon is becoming the most popular and, in many settings, preferred modality for CRC screening in the general population. As far back as the early 1990s, sigmoidoscopy was shown to reduce CRC incidence and mortality. Just as technological advances since the 1960s advanced OC to the forefront of CRC screening, more recent technological advances have introduced alternative methods of polyp detection.

CTC has taken root as a viable alternative to OC and has comparable detection rates for large polyps. A recent American College of Radiology white paper on the status of CTC in 2009 summarizes in detail the various features of CTC related to screening for CRC.[64] Debate continues regarding the miss rate of both modalities, primarily for lesions <9 mm. The clinical course of such lesions is not well described, and this further complicates the issue. The long-term effects of radiation, the possible lack of generalizability of the study populations, and the management of extracolonic findings have also raised concerns. For these reasons, in May 2009, the Centers for Medicare and Medicaid Services passed the following national noncoverage decision:

> The evidence is inadequate to conclude that CTC is an appropriate CRC screening test under § 1861(pp)(1) of the Social Security Act. CTC for CRC screening remains noncovered.

Before CTC can be implemented in the population as a CRC screening tool, quality metrics will need to be defined and placed into clinical practice. Variability in reader training and experience has been cited as a major cause in the discrepancy in CTC accuracy in nonscreening populations.[12] More information is needed with respect to patient preferences for CRC screening tests. If our goal is to get more people screened for CRC, this should be considered if some patients would prefer CTC over OC. Also, because most screening OC examinations are normal, CTC could be a helpful prescreen, with only abnormal CTC examinations triaged for OC, and this might also be theoretically cost-effective.

In conclusion, there is much room for improvement in the screening rates for CRC. The development of less invasive, more widely acceptable techniques for screening, such as CTC, deserves attention as a way of accomplishing this goal. Current data suggest that OC will remain the gold standard for CRC screening in the near future. Physicians should remain alert to developing data in the field to decide the best path for their patients.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. *CA Cancer J Clin* 2008; 58: 71–96.
2. Jemal A, Clegg L, Ward E, et al. Annual report to the nation on the status of cancer, 1975–2001, with a special feature regarding survival. *Cancer* 2004; 101: 3–27.
3. Centers for Disease Control and Prevention. Colorectal cancer screening rates. http://www.cdc.gov/cancer/colorectal/statistics/screening_rates.htm. Accessed January 2010.
4. Schenck AP, Peacock SC, Klabunde CN, et al. Trends in colorectal cancer test use in the Medicare population. *Am J Prev Med* 2009; 37: 1–7.
5. Joseph DA, Rim SH, Seeff LC. Use of colorectal cancer tests–United States, 2002, 2004, and 2006. *MMWR Morb Mortal Wkly Rep* 2008; 57: 253–258.
6. Levin B, Lieberman D, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology* 2008; 134: 1570–1595.
7. US Preventive Services Task Force. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2008; 149: 627–637.
8. Rex DK, Johnson DA, Anderson JC, et al. American College of Gastroenterology guidelines for colorectal cancer screening 2009 [corrected]. *Am J Gastroenterol* 2009; 104: 739–750.
9. New York City Department of Health and Mental Hygiene. Colon cancer. http://www.nyc.gov/html/doh/html/cancer/cancercolonscreen.shtml. Accessed January 2010.
10. Vining DJ, Gelfand DW, Bechtold RE, et al. Technical feasibility of colon imaging with helical CT and virtual reality [abstract]. *AJR Am J Roentgenol* 1994; 162(suppl): 104.

11. Burling D, Hallingan S, Slater A, et al. Potentially serious adverse events at CT colonography in symptomatic patients: national survey of the United Kingdom. *Radiology* 2006; 239: 464–471.
12. Pickhardt PJ, Choi JR, Hwang I, et al. Computer tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N Engl J Med* 2003; 349: 2191–2200.
13. Whitlock EP, Lin JS, Liles E, et al. Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2008; 149: 638–658.
14. Bond JH. Clinical evidence for the adenoma-carcinoma sequence, and the management of patients with colorectal adenomas. *Semin Gastrointest Dis* 2000; 11: 176–184.
15. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med* 1993; 329: 1977–1981.
16. Citarda F, Tomaselli G, Capocaccia R, et al. Efficacy in standard clinical practice of colonoscopic polypectomy in reducing colorectal cancer incidence. *Gut* 2001; 48: 812–815.
17. Zauber A, Winawer SJ, O'Brien MJ, et al. Significant long term reduction in colorectal cancer mortality with colonoscopic polypectomy: findings of the National Polyp Study. *Gastroenterology* 2007; 132: A50.
18. Rex DK, Eid E. Considerations regarding the present and future roles of colonoscopy in colorectal cancer prevention. *Clin Gastroenterol Hepatol* 2008; 6: 506–514.
19. Zauber A, Winawer SJ, Lansdorp-Vogelaar I, et al. Effect of initial polypectomy versus surveillance polypectomy on colorectal cancer mortality reduction: micro-simulation modeling of the National Polyp Study. *Am J Gastroenterol* 2007; 102: S558.
20. Kahi CJ, Imperiale TF, Juliar BE, Rex DK. Effect of screening colonoscopy on colorectal cancer incidence and mortality. *Clin Gastroenterol Hepatol* 2009; 7: 770–775.
21. Singh H, Turner D, Xue L, et al. Risk of developing colorectal cancer following a negative colonoscopy examination: evidence for a 10-year interval between colonoscopies. *JAMA* 2006; 295: 2366–2373.
22. Baxter N, Goldwasser M, Paszat L, et al. Association of colonoscopy and death from colorectal cancer. *Ann Intern Med* 2009; 150: 1–8.
23. Rex DK, Cutler CS, Lemmel GT, et al. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology* 1997; 112: 24–28.
24. Imperiale T, Glowinski E, Lin-Cooper C, et al. Five-year risk of colorectal neoplasia after negative screening colonoscopy. *N Engl J Med* 2008; 359: 1218–1224.
25. Rex D, Cummings O, Helper D, et al. 5-year incidence of adenomas after negative colonoscopy in asymptomatic average-risk persons. *Gastroenterology* 1996; 111: 1178–1181.
26. van Rijn J, Reitsma J, Stoker J, et al. Polyp miss rate determined by tandem colonoscopy: a systematic review. *Am J Gastroenterol* 2006; 101: 343–350.
27. Pickhardt PJ, Nugent PA, Mysliwiec PA, et al. Location of adenomas missed by optical colonoscopy. *Ann Intern Med* 2004; 141: 352–359.
28. Winawer SJ, Zauber AG. The advanced adenoma as the primary target of screening. *Gastrointest Endosc Clin N Am* 2002; 12: 1–9.
29. Fletcher R. Colorectal cancer screening on stronger footing. *N Engl J Med* 2008; 359: 1285–1287.
30. Goldstein NS. Serrated pathway and APC (conventional)-type colorectal polyps: molecular-morphologic correlations, genetic pathways, and implications for classification. *Am J Clin Pathol* 2006; 125: 146–153.
31. Saitoh Y, Waxman I, West A, et al. Prevalence and distinctive biologic features of flat colorectal adenomas in a North American population. *Gastroenterology* 2001; 120: 1657–1665.
32. Fujii T, Rembacken BJ, Dixon MF, et al. Flat adenomas in the United Kingdom: are treatable cancers being missed? *Endoscopy* 1998; 30: 437–443.
33. Soetikno R, Kaltenbach T, Rouse R, et al. Prevalence of nonpolypoid (flat and depressed) colorectal neoplasms in asymptomatic and symptomatic adults. *JAMA* 2008; 299: 1027–1035.
34. Hurlstone DP, Brown S, Cross SS. The role of flat and depressed colorectal lesions in colorectal carcinogenesis: new insights from clinicopathological findings in high-magnification chromoscopic colonoscopy. *Histopathology* 2003; 43: 413–426.
35. O'Brien MJ, Winawer SJ, Zauber AG, et al. Flat adenomas in the National Polyp Study: is there increased risk for high-grade dysplasia initially or during surveillance? *Clin Gastroenterol Hepatol* 2004; 2: 905–911.
36. Rex DK, Imperiale TF, Latinovich DR, Bratcher LL. Impact of bowel preparation on efficiency and cost of colonoscopy. *Am J Gastroenterol* 2002; 97: 1696–1700.
37. Barclay RL, Vicari JJ, Doughty AS, et al. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. *N Engl J Med* 2006; 355: 2533–2541.
38. Simmons DT, Harewood GC, Baron TH, et al. Impact of endoscopist withdrawal speed on polyp yield: implication for optimal colonoscopy withdrawal time. *Aliment Pharmacol Ther* 2006; 24: 965–971.
39. Adler A, Aschenbeck J, Yenerim T, et al. Narrow-band versus white-light high definition television endoscopic imaging for screening colonoscopy: a prospective randomized trial. *Gastroenterology* 2009; 136: 410–416.
40. East JE, Stavrindis M, Thomas-Gibson S, et al. A comparative study of standard vs. high-definition colonoscopy for adenoma and hyperplastic polyp detection with optimized withdrawal technique. *Aliment Pharmacol Ther* 2008; 28: 768–776.
41. Kaltenbach T, Friedland S, Soetniko R. A randomised tandem colonoscopy trial of narrow band imaging versus white light examination to compare neoplasia miss rates. *Gut* 2008; 57: 1406–1412.
42. Rastogi A, Bansal A, Wani S, et al. Narrow-band imaging colonoscopy–a pilot feasibility study for the detection of polyps and correlation of surface patterns with polyp histologic diagnosis. *Gastrointest Endosc* 2008; 67: 280–286.
43. Matsuda T, Saito Y, Fu K, et al. Does autofluorescence imaging videoendoscopy system improve the colonoscopic polyp detection rate?–A pilot study. *Am J Gastroenterol* 2008; 103: 1926–1932.
44. Brooker JC, Saunders BP, Shah SG, et al. Total colonoscopic dye-spray increases the detection of

diminutive adenomas during routine colonoscopy: a randomized controlled trial. *Gastrointest Endosc* 2002; 56: 333–338.

45. Rabeneck L, Paszat L, Hilsden R, et al. Bleeding and perforation after outpatient colonoscopy and their risk factors in usual clinical practice. *Gastroenterology* 2008; 135: 1899–1906.
46. Gatto NM, Frucht H, Sundararajan V, et al. Risk of perforation after colonoscopy and sigmoidoscopy: a population-based study. *J Natl Cancer Inst* 2003; 95: 230–236.
47. Newcomb P, Norfleet RG, Storer B, et al. Screening sigmoidoscopy and colorectal cancer mortality. *J Natl Cancer Inst* 1992; 84: 1572–1575.
48. Selby JV, Friedman GD, Quesenberry CP Jr, Weiss NS. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med* 1992; 326: 653–657.
49. Winawer SJ, Zauber AG, O'Brien MJ, et al. Randomized comparison of surveillance intervals after colonoscopic removal of newly diagnosed adenomatous polyps. *N Engl J Med* 1993; 328: 901–906.
50. Neugut AI, Jacobson JS, Ahsan H, et al. Incidence and recurrence rates of colorectal adenomas: a prospective study. *Gastroenterology* 1995; 108: 402–408.
51. Squillance S, Berggreen P, Jaffe P, et al. A normal initial colonoscopy after age 50 does not predict a polyp-free status for life. *Am J Gastroenterol* 1994; 89: 1156–1159.
52. Martinez ME, Baron JA, Lieberman DA, et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. *Gastroenterology* 2009; 136: 832–841.
53. Mulhall BP, Veerappan GR, Jackson JF. Meta-analysis: computed tomographic colonography. *Ann Intern Med* 2005; 142: 635–650.
54. Kim DH, Pickhardt PJ, Taylor AJ, et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. *N Engl J Med* 2007; 357: 1403–1412.
55. Schoenfeld P, Cash B, Flood A, et al. Colonoscopic screening of average-risk women for colorectal neoplasia. *N Engl J Med* 2005; 352: 2061–2068.
56. Imperiale TF, Wagner DR, Lin CY, et al. Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. *N Engl J Med* 2000; 343: 169–174.
57. Regula J, Rupinski M, Kraszewska E, et al. Colonoscopy in colorectal-cancer screening for the detection of advanced neoplasia. *N Engl J Med* 2006; 355: 1863–1872.
58. Johnson CD, Chen MH, Toledano AY, et al. Accuracy of CT colonography for detection of large adenomas and cancers. *N Engl J Med* 2008; 359: 1207–1217.
59. Summerton S, Little E, Cappell MS. CT colonography: current status and future promise. *Gastroenterol Clin N Am* 2008; 37: 161–189.
60. Brenner DJ, Georgsson MA. Mass screening with CT colonography: should the radiation exposure be of concern? *Gastroenterology* 2005; 129: 328–337.
61. Sosna J, Blackhar A, Amitai M, et al. Colonic perforation at CT colonography: assessment of risk in a multicenter large cohort. *Radiology* 2006; 239: 457–463.
62. Lieberman D, Moravec M, Holub J, et al. Polyp size and advanced histology in patients undergoing colonoscopy screening: implications for CT colonography. *Gastroenterology* 2008; 135: 1100–1105.
63. Pickhardt PJ, Hassan C, Laghi A, et al. Clinical management of small (6- to 9-mm) polyps detected at screening CT colonography: a cost-effectiveness analysis. *AJR Am J Roentgenol* 2008; 191: 1509–1516.
64. McFarland EG, Fletcher JG, Pickhardt P, et al. ACR Colon Cancer Committee white paper: status of CT colonography 2009. *J Am Coll Radiol* 2009; 6: 756–772.

Mount Sinai Journal of Medicine 77:225–235, 2010

SPECIAL FEATURE

Epigenetics in Women's Health Care

Yevgeniya Pozharny, MD, Luca Lambertini, PhD, Garfield Clunie, MD,
Lauren Ferrara, MD, and Men-Jean Lee, MD

Mount Sinai School of Medicine, New York, NY

OUTLINE

Epigenetic Mechanisms
  DNA Methylation
  Histone Modification
  RNA Regulation
Epigenetics and the Life Cycle
  Fertilization
  Early Embryo Development
  Gametogenesis
  Age-Related Reprogramming
Epigenetics and Obstetrics
  Genomic Imprinting and Development
  Fetal Programming
  Intrauterine Growth Restriction
  Epigenetics and Preeclampsia
Epigenetics and Diet
  Dietary Methyl Donors and Cofactors
  High Fat or Glucose Intake
Environmental Factors
  Heavy Metals
  Endocrine Disruptors
Conclusion

ABSTRACT

Epigenetics refers to structural modifications to genes that do not change the nucleotide sequence itself but instead control and regulate gene expression. DNA methylation, histone modification, and RNA regulation are some of the mechanisms involved in epigenetic modification. Epigenetic changes are believed to be a result of changes in an organism's environment that result in fixed and permanent changes in most differentiated cells. Some environmental changes that have been linked to epigenetic changes include starvation, folic acid, and various chemical exposures. There are periods in an organism's life cycle in which the organism is particularly susceptible to epigenetic influences; these include fertilization, gametogenesis, and early embryo development. These are also windows of opportunity for interventions during the reproductive life cycle of women to improve maternal-child health. New data suggest that epigenetic influences might be involved in the regulation of fetal development and the pathophysiology of adult diseases such as cancer, diabetes, obesity, and neurodevelopmental disorders. Various epigenetic mechanisms may also be involved in the pathogenesis of preeclampsia and intrauterine growth restriction. Additionally, environmental exposures are being held responsible for causing epigenetic changes that lead to a disease process. Exposure to heavy metals, bioflavinoids, and endocrine disruptors, such as bisphenol A and phthalates, has been shown to affect the epigenetic memory of an organism. Their long-term effects are unclear at this point, but many ongoing studies are attempting to elucidate the pathophysiological effects of such gene-environment interactions. ***Mt Sinai J Med 77:225–235, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: epigenetics, imprinted genes, placenta.

Address Correspondence to:

Men-Jean Lee
Mount Sinai School of Medicine
New York, NY
Email: menjean.lee@mssm.edu

The basis of cell differentiation lies in the ability of tissues to express different genes from the same basic set of genetic information stored in DNA. The human genome contains approximately 30,000 genes, and yet only a small number of them are turned on or transcribed within a particular tissue type. This small subset of genes allows the cells to produce proteins unique to their functions. Gene expression is controlled by mechanisms that act at the DNA level and by posttranscriptional and posttranslational modifications.[1] Because these

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20176

© 2010 Mount Sinai School of Medicine

processes do not change the genotype by directly mutating the DNA sequence, they are termed *epigenetic modifications*.[2,3]

However, the definition of *epigenetics* is broad and nonspecific. Initially, *epigenetics* was used as a term that defined ways in which the environment influences phenotype (the way that an organism looks). This definition was proposed by C. H. Waddington, who observed that environmental factors can alter the expected phenotype to produce an entirely new one from the same genotype.[4] Waddington also noted that this new phenotype can become heritable even in the absence of the original environmental stimulus. His experiments on *Drosophila melanogaster* demonstrated this concept. Waddington showed that the vein pattern on the wing produced by heat shock can lead to a stable population of flies exhibiting the new phenotype even in the absence of heat shock and without direct gene mutation in the subsequent generations.[1]

Today, *epigenetics* is used to describe functional modifications to genes that do not change the nucleotide sequence itself. *Epigenetic inheritance*, in turn, refers to biological processes that govern heritable changes in gene expression without altering the basic DNA sequence. In most tissues in the body, these epigenetic changes become fixed once the cell differentiates or stops participating in the cell cycle.[1] However, in some developmental stages and certain disease states, some cells can go through epigenetic reprogramming.[2,3,5] Such epigenetic changes that occur during these specific situations (eg, pregnancy and aging) are the focus of this discussion.

EPIGENETIC MECHANISMS

There are 3 basic epigenetic systems that control and regulate gene expression: DNA methylation, histone modification, and RNA regulation. These independent mechanisms interact to modulate the response of the cell to environmental stimuli. When their setting is modified during fetal development, they become heritable because of the substantial change that is imposed on the developing primordial germinal cells.

DNA Methylation

This was the first epigenetic mechanism that was described in the literature. It has been widely explored in its significance and function. *DNA methylation* refers to the addition of a methyl group ($-CH_3$) to the cytosine nucleotides as a part of the CpG dinucleotide sequence. It is generally accepted that the DNA methylation status correlates with the activity of a specific genomic region. Methylated DNA is normally found in areas that are inactive or silent (they do not get transcribed), whereas unmethylated regions correspond to actively transcribed DNA regions.[6,7] It is thought that 3% of the cytosines in the human genome are methylated. Although the overall occurrence of CpG dinucleotides in the human genome is significantly less than expected according to the predicted frequency, there are also stretches of DNA that are rich in CpG dinucleotides and extend for hundreds of bases. These are called CpG islands.[1,6] However, their significance is still being debated. Although it was originally thought that the CpG islands represent or include regulatory regions, it has since been shown that the most of them do not seem to participate in any regulatory pathways. Therefore, the role played by DNA methylation strictly correlates with the DNA area of placement in which inactive genomic areas, including CpG islands, become methylated and demethylated as a result of the general cellular activity.

Enzymes responsible for DNA methylation are known as DNA methyltransferases (DNMTs). There are 2 types of DNMTs. DNMT3a and DNMT3b are responsible for DNA methylation at previously unmethylated sites. These enzymes are the methylases. DNMT1, on the other hand, maintains methylation at preexisting sites (maintenance methylase). The preference of DNMT1 for hemimethylated sites is what allows the organism to maintain heritable patterns of methylation. In mammalian semiconservative DNA replication, this enzyme recognizes the methylation of the parental strand and matches it to the new daughter strand of DNA.[1–3] The exposure to environmental conditions that specifically alter the original DNA methylation signal during the early phases of fetal development can therefore become heritable and produce phenotypic modifications despite the lack of alteration to the DNA sequence itself.

Histone Modification

In a human cell, DNA is usually wrapped around the histone proteins and thus forms chromatin (a nucleoprotein complex).[8] The structure of chromatin is vital to gene expression because it determines accessibility to the underlying DNA and the recruitment of the transcriptional factors and other regulatory molecules. Changes in chromatin structure and compaction are achieved through chemical modifications of the amino acids, which make up the histone protein backbone. Histones are modified through methylation, acetylation, phosphorylation, ubiquitination, and deoxyadenosine-5'-diphosphate (dADP) ribosylation.[9–11]

For example, histone acetylation is thought to be associated with relaxation of chromatin and, therefore, allows increased access to gene regulatory proteins, including transcription and elongation factors. The addition of acetyl groups on lysine residues of core histones in the nucleosome decreases their binding to DNA, and this leads to a more open chromatin structure and increased access to transcription factors. In other words, histone acetylation promotes gene transcription. On the other hand, transciptionally inactive chromatin is associated with histone methylation because methylation prohibits binding of transcription factors. In general, histone methylation inhibits gene expression. Hence, a complex mechanism of epigenetic modification of DNA itself as well as the associated histones determines transcriptional regulation of a given gene.[8]

There is evidence to suggest that gene expression is affected by the genomic context of the genes with respect to their chromosomal location. Each chromosome is organized into functional domains and silent domains. For instance, when genes are transolcated onto a new chromosomal region (in both spontaneous breakage and iatrogenic manipulations), these genes are expressed in an aberrant way, even though all the required control sequences are preserved.[12]

RNA Regulation

RNA regulation, such as RNA silencing, represents the more complex and least known epigenetic regulation mechanism. Its importance was first proposed when researchers began to explore the plethora of non–protein-coding RNAs (ncRNAs) coded by the human genome. Humans transcribe roughly as many protein-coding RNA species as the worm *Caenorhabditis elegans*, a model widely used for the study of molecular and developmental biology; however, although *C. elegans* does not code for any ncRNA, *Homo sapiens* is supposed to transcribe several hundred thousands of ncRNAs. This peculiar fact has led several scientists to hypothesize that this phenomenon is at the basis of the evolution of the species. ncRNAs are classified into several groups based on their length and function, which invariably regulate and fine-tune gene expression. Briefly, micro-RNAs, 20 to 22 base pairs long, down-regulate gene expression via complementary binding of the 3′ untranslated region of target RNAs; short interfering RNAs, 21 to 25 base pairs long, lead to RNA degradation via specific binding and operate in posttranslational gene silencing; small nuclear RNAs, 100 to 300 base pairs long, are involved in RNA splicing and telomere maintenance; small nucleolar RNAs of variable length are involved in ribosomal RNA processing; and long ncRNAs include tens of thousands of transcripts that vary from a few base pairs to many kilobase pairs in length. Their transcription is often critical for proper cell function through pathways that are unknown at this time. Recently, it has been proposed that long ncRNAs play a critical role in the setting and maintenance of the genomic imprinting signal by regulating access to the imprinting control regions during fetal development and in adult life.

EPIGENETICS AND THE LIFE CYCLE

Epigenetic marks are fixed for most fully differentiated cell types. They reflect the genotype, developmental history, and environmental influences of that particular cell type. As described earlier, certain developmental stages and disease states are associated with epigenetic reprogramming.[2,13] Removal of old epigenetic marks is followed by the creation of a new set of marks. Therefore, these times of intense epigenetic changes can have significant effects on the life of an individual as well as his or her progeny.

Most of our understanding of the changes in the epigenetic code comes from animal models (primarily mice). However, this information easily extends to humans. Epigenetic reprogramming can be observed in 4 distinct phases: fertilization, early embryo development, gametogenesis, and age-related reprogramming.

Fertilization

DNA methylation and histone modification signals are established during fetal development with a carefully planned schedule of events. The DNA methylation profile undergoes a reprogramming phase that takes place during the early phases of embryonic development. Epigenetic reprogramming acts bimodally during the preimplantation period and differentially affects the embryo and the primordial phase of gametogenesis. During the preimplantation period, the embryo undergoes genome-wide demethylation. Both parental genomes are remethylated around the time of implantation. Concomitantly, methylated imprinted genes are protected from the demethylation wave.

Active demethylation is the first phase of epigenetic reprogramming that occurs between fertilization and formation of the blastocyst. It involves rapid paternal-specific loss of methylation. Although this process involves the entire paternal genome, it spares paternally imprinted genes, heterochromatin, and some repetitive elements.[13] The mechanism for this rapid process is not well understood, although it may involve the nucleus of the sperm and cytoplasmic maturity of the oocyte.[14,15]

Active demethylation is followed by passive demethylation. Passive demethylation is a process that occurs because of the absence of maintenance methylation during DNA replication.[16] Therefore, new DNA strands do not become methylated, and this leads to the overall decrease in the methyl content.

Histone programming relies on asymmetric chromatin remodeling, which is required for the completion of fertilization. At fertilization, the oocyte chromatin is organized into DNA complexes with histones, whereas sperm chromatin carries protamines instead of histones. An exchange of nucleoprotamine for nucleohistone occurs during the first hour after fertilization, and this is concurrent with the beginning of the first DNA demethylation wave. This results in an asymmetric distribution of chromatin in the fertilized oocyte. Paternal chromatin then undergoes a removal of acetylation groups catalyzed by histone deacetylases, which is followed by the recruitment of specific histone methyltransferases that also contribute to the placement of imprinting signals. An inability to complete this transition may contribute to the loss of DNA methylation specifically from the male pronucleus. At syngamy, maternal and paternal chromosomes align and maintain their parental associations. This sequence of events highlights epigenetic differences between maternal and paternal chromosomes arising during zygote programming that may very well remain throughout the preimplantation period and influence the proper programming of genomic imprinting.[17]

Early Embryo Development

The first differentiation event, which establishes 2 cell lineages in the embryo, overlaps with the start of methylation in the embryo. The 2 lineages are the inner cell mass, which forms all the tissues of the adult, and the trophectoderm, which eventually produces placental structures. It turns out that the inner cell mass becomes hypermethylated, whereas the trophectoderm becomes hypomethylated.[18,19] This methylation pattern is thought to be preserved throughout pregnancy.

Gametogenesis

Primordial germ cells are derived from the inner cell mass of the developing embryo. During early development, these cells appear to be highly methylated; however, methylation is no longer noted as these cells migrate to the genital ridge. Later in fetal development, they migrate to the developing germinal ridges and become mature gametes.[20–22] After the first methylation wave, these cells undergo the second wave targeted at imprinted domains. It erases the parental contribution in primordial germ cells. It appears that this loss of methylation is regulated by an active targeted process of DNA demethylation.[23] Maximal demethylation occurs around 12 to 13 days after fertilization.[22] Imprinting marks are then replaced with new signals consistent with the gender of the developing embryo; at that point, gametogenesis is halted.

This unique setting clearly defines a window of vulnerability, particularly during the early phases of fetal development, that extends into the later phases, and it exerts a transgenerational effect by perturbing the primordial germ cells.

Age-Related Reprogramming

Epigenetic patterns are not fixed for life. Studies indicate that global hypomethylation occurs in aging humans. In other words, demethylation occurs over time. This process is tissue-specific and affects gene promoters and noncoding DNA sequences. It has been demonstrated to occur in the brain, liver, small intestine mucosa, spleen, heart, and T lymphocytes.[17,24,25]

Interestingly, age-related diseases such as cancers and autoimmune disorders have been shown to be associated with global hypomethylation of DNA, which leads to increased gene transcription. On the other hand, hypermethylation of tumor suppressor gene CpG islands also increases the risk of cancer by inhibiting DNA transcription of tumor suppressors and thus allowing procancerous genes to be transcribed. Data show a linear increase in the methylation of the promoter region of various cancer suppressor genes with age.[17,24–28]

No definite mechanism has been proposed to explain changes in methylation patterns in the aging population. Some studies have associated these changes with exposure to DNA methylation inhibitors, such as drugs or diet, whereas others have suggested that methylation declines because of a decrease in the effective methyltransferase (DNMT) expression.

EPIGENETICS AND OBSTETRICS

Characterized by high rates of DNA synthesis, the early phases of embryogenesis represent the critical period of fetal programming when most of the epigenetic signals necessary for normal development are established.[29] The proper establishment of genomic imprinting patterns is fundamental during this time period because it can substantially affect fetal development and the well-being of the future adult organism. These events are thought to have been developed to be responsive to the influence of intrauterine environmental factors in order to prepare the newborn to face environmental conditions. Although reversible, epigenetic marks may persist through the generations despite the absence of continued exposure in progeny.[29–33]

Genomic Imprinting and Development

Genomic imprinting refers to the silencing of 1 parental allele in the fertilized zygote. This process is driven by DNA methylation, histone modification, and RNA regulation and leads to monoallelic expression of these genes in the offspring. This process results in reversible, parent of origin–specific marks in the genome that produce the functional differences in the genetic information contributed by each parent.[34,35] Imprinted genes are functionally haploid, so genomic imprinting is counter to classical Mendelian genetic theory (equal inheritance of parental traits). In humans, the estimated number of imprinted genes is around 1% of the genome or approximately 200 genes. The exact number of human imprinted genes is difficult to determine, however, because monoallelic expression of imprinted genes may occur in 1 specific isoform, only in specific tissue types, or only at a specific stage of development. At the present time, 87 genes have been experimentally shown to be imprinted in at least 1 organ/tissue, whereas some 100 additional genes have been computationally predicted to be imprinted.

The contribution of genomic imprinting to placental and fetal development has been explored in many studies, which have primarily used animal models. Its fundamental role has been demonstrated unequivocally by a recently study in mice. Bimaternal embryos were generated with oocytes with no epigenetic imprinting marks established. These embryos, carrying complete sex-specific imprinting signaling, were not viable. In fact, they were severely growth-restricted, whereas their placentas were small with abnormal histology. Deletion of 2 genomic areas corresponding to 2 imprinted gene clusters [insulin-growth factor 2-H19 (IGF2-H19) and delta-like 1 homolog (Drosophila)–deiodinase iodothyronine type III (*dlk1-dio3*)] in the unmarked oocytes produced normal placental formation with normal growth and development of the fetus into adulthood. These features underline the great importance of imprinted genes to promoting the proper development of the placenta and the fetus.

Perturbations of the genomic imprinting status, known as *loss of imprinting*, result in the reactivation of the silent allele leading to biallelic expression of the affected gene(s), which may increase gene dosage. Loss of imprinting has been associated with several developmental disorders, including (1) diseases with parent-of-origin effect such as Angelman, Prader-Willi, Beckwith-Wiedemann, and Silver-Russell syndromes,[36] Huntington's disease, cystic fibrosis, and spina bifida; (2) neurological disorders such as Rett syndrome, autism spectrum disorders, bipolar disorder, Tourette syndrome, and schizophrenia[37]; (3) reproductive disorders such as intrauterine growth restriction (IUGR) and preeclampsia; and (4) cancers such as hepatoblastoma, colorectal and bladder cancer, lymphoblastic leukemia, Wilms' tumor, and ovarian and breast cancer.

Fetal Programming

The Barker hypothesis refers to the idea that adult disease originates during fetal life. There are many data supporting the epidemiological observation that individuals who were smaller at birth are at increased risk for developing coronary heart disease, stroke, type 2 diabetes mellitus, obesity, metabolic syndrome, and osteoporosis as adults.[38–42] This observation indicates that developmental plasticity, which has been defined as "the ability of an organism to change irreversibly its structure and function during a critical time window in response to an environmental stimulus," exists.[43,44] The fetus appears to develop adaptive responses to the intrauterine environment through cues that it receives from its mother regarding her physical state. These responses may manifest as changes in metabolism, hormone production, and tissue sensitivity to hormones and lead to persistent alterations in homeostatic set points.

Interestingly, experimental models suggest that environmental challenges during pregnancy lead to promoter methylation. This can alter gene expression in multiple pathways involved in physiological processes. For instance, in rat models, a protein-restricted diet alters promoter methylation and leads to changes in the hepatic glucocorticoid receptor and the peroxisome proliferator-activated receptor (PPAR).[45,46] These alterations affect carbohydrate and lipid metabolism via increased expression of gluconeogenic enzymes and increased blood glucose levels.[47] Also, similar epigenetic changes have been noted in p53 in the kidneys and in the angiotensin II type 1b receptor in the adrenal gland; these, in turn, influence renal apoptosis and pressor responses in the hypothalamic glucocorticoid receptor,[48] which influences an organism's response to stress.

Phenotypic expression of epigenetic modifications during development may not become obvious until later in life. This is particularly true in situations in which the genes that modify responses to later environmental factors (eg, a high lipid diet) are affected. It is thought that epigenetic changes usually occur from the periconceptional period into postnatal life,[49,50] although the true time window is unknown. Evidence stemming from twin studies suggests that some of the epigenetic changes can occur later in life and have a relationship with age and environmental exposure.[51]

Epigenetic changes can be reversed upon exposure to new environmental influences. As described previously, PPAR-alpha is a gene that is associated with diabetes and obesity. Exogenous leptin administration to the neonatal offspring of malnourished rats has been shown to correct phenotypic changes, gene expression, and associated methylation changes in PPAR-alpha.[52,53] Dietary intervention with *n*-3 fatty acids seems to reverse hyperleptinemia and hypertension.[54] Maternal folic acid supplementation of protein-deprived rats seems to prevent glucose intolerance and increased expression of gluconeogenic enzymes in the offspring.[47] These promising findings seem to suggest that ultimately susceptible individuals might be identified via screening for epigenetic markers in early life, and customized dietary and environmental changes can be instituted.

Intrauterine Growth Restriction

Although it is generally accepted that IUGR is associated with a limited nutrient and oxygen supply, the specific mechanisms involved in IUGR are mostly unknown. Fetal growth is largely determined by the nutrient supply, which is dependent on placental transport functions. Net placental transport of nutrients is multifactorial. It depends on the concentration gradient between maternal and fetal blood, the placental blood flow and metabolism, and the expression and activity of specific membrane-bound transporters. Studies show that blood flow is decreased on both sides of the placental exchange barrier in pregnancies complicated by IUGR.[55,56] However, a decrease in blood flow alone cannot explain decreased levels of nutrients seen in IUGR.

Evidence suggests that maternal undernutrition during pregnancy reduces placental and fetal growth in domesticated animals and humans.[57,58] It appears that the peri-implantation period and period of rapid placental development are times during which fetal growth is most susceptible to maternal nutritional deficiency.[25] For instance, maternal dietary deficiency in amino acids leads to a global reduction of genomic methylation.[59]

Regardless of what the exact mechanism for IUGR might be, 1 pathway most likely involves the ratio of growth hormone to IGF and direct changes in the growth factors and their interaction with their receptors.[60] For example, the IGF2–H19 gene complex is imprinted, and IGF2 is paternally expressed in both the fetus and the placenta. This gene cluster plays a role in matching the placental nutrient supply to fetal nutrient demand. In this case, differential methylation affects imprinting. Both an increase and a decrease in IGF2 activity play a role in placental size and efficiency.[61] It appears that early dietary deficiency may affect methylation patterns and, therefore, change placental function and fetal development. In fact, when this gene is deleted in mice, passive permeability for nutrients in the placenta is decreased, whereas secondary active transport of amino acids is up-regulated. Therefore, compensation for the decrease in passive permeability occurs.[62,63] IUGR develops when this compensatory mechanism fails later in the pregnancy.

Epigenetics and Preeclampsia

Preeclampsia is commonly considered a 2-stage disease.[64] The first part of the pathophysiology occurs when early trophoblast invasion and remodeling of the spiral arteries fail. Consequently, blood flow to the placenta becomes limited, and a hypoxic environment ensues. As a result, oxidative stress induces ischemic lesions.[65,66] The maternal syndrome of preeclampsia, as demonstrated by an exaggerated systemic inflammatory response, is the second stage of the process.[67] The inflammatory response involves leukocyte and endothelial dysfunction, thrombosis, and rennin-angiotensin-aldosterone system activation.[68,69] It appears that the stressed placenta leads to clinical manifestations of preeclampsia.

Epigenetics may possibly be involved in the development of preeclampsia. It has been well established that the risk of preeclampsia is increased in those pregnancies conceived with the help of assisted reproductive technology, and this implies that early epigenetic modification of the gametes and/or embryo may be sufficient to cause placental defects.[70] We also know that deletion of a certain imprinted gene (p57-Kip2) in a mouse model can lead to hypertension and proteinuria during pregnancy, which also happen to be hallmark manifestations of preeclampsia in humans.[71] Further studies in mice indicate that the development of preeclampsia is dependent on the environment.[72]

Epigenetic alterations to nonimprinted genes also occur in preeclampsia. For instance, most recently, the serine protease (SERPIN) family of genes has come under investigation as a family of genes induced by hypoxia and preeclampsia. The SERPIN genes encode at least 36 proteins with conserved structure but various functions.[1] These proteins act as protease inhibitors, storage proteins (ovalbumin), carrier proteins (steroid-binding globulins), and hormone precursors (angiotensinogen). In short, all these proteins are involved in the maintenance of homeostasis through regulation of inflammation, coagulation, fibrinolysis, complement activation, and phagocytosis.

Recent data suggest that methylation changes of the SERPIN5 and SERPIN3 gene promoters have been demonstrated in pregnancies affected with preeclampsia.[73] It appears that there is significant hypomethylation in the promoters of both genes in placentas of women with preeclampsia in comparison with control placentas. Understanding the epigenetic markers that put a woman at risk for developing preeclampsia may enable us to use this knowledge in clinical practice. It may be possible to look for these epigenetic markers in the samples obtained during chorionic villus sampling and in fetal DNA found in the maternal blood stream, and this could allow for the early prediction of the development of preeclampsia in an individual patient.

EPIGENETICS AND DIET

Dietary Methyl Donors and Cofactors

The methyl groups used in DNA methylation come from 2 sources: synthesis and diet. Therefore, diet has a significant impact on global methylation patterns. The daily diet of humans consists of approximately 60 mmol of methyl groups. Both excesses and deficiencies of many of the dietary elements alter methylation patterns.

Folate is a valuable cofactor in methyl metabolism. It is used to produce *S*-adenosyl methionine, which is the principal methyl donor. It has been shown that folate depletion alone is sufficient to diminish the methyl pool.[74] Folate intake is the sum of dietary folate and supplemental folate. Folate found in foods is predominantly in the form of 5-methyl-tetrahydrofolate. Fully unreduced (folic acid) and partially reduced (dihydrofolate) forms are seen as well. In contrast, folate found in supplements is in its synthetic form, folic acid, which needs to be reduced prior to its involvement in cellular metabolism.[75]

It is not surprising that changes in the folate supply affect DNA methylation. For instance, in pregnancy, low levels of circulating folate are associated with an increased risk of preterm delivery,[76] IUGR,[77] poor growth of the placenta and uterus,[77] and placental infarctions and premature rupture of membranes.[78] In turn, elevated homocysteine levels have been linked to recurrent miscarriage,[79] placental abruption,[80,81] and preeclampsia.[82] It is likely that epigenetic mechanisms are involved. Data in mice indicate that changes in early nutrition can alter epigenetic patterns in a young embryo and therefore affect all tissues.

However, it is unclear whether synthetic folate and dietary folate are equivalent in their ability to decrease complications associated with pregnancy and reproduction. One area in which folic acid supplementation has been studied extensively is the prevention of neural tube defects. Laurence *et al.*[83] and Smithells *et al.*[84] were among the first to show that supplementation with folic acid (synthetic folate) is effective at preventing neural tube defects. Their findings have been supported by many others. In fact, in 1991, the Centers for Disease Control and Prevention recommended that all women with a previous history of a fetus with neural tube defects take supplemental folic acid. Following that recommendation, the US Public Health Service advised all women of child-bearing age to ingest supplemental folic acid.

It is also unclear whether there is a strict linear relationship between folate intake and a reduction in pregnancy complications. For instance, current literature suggests that decreased folate intake is associated with an increased risk of breast cancer. Additionally, there is evidence that adequate folate intake can mitigate the increased risk associated with alcohol consumption.[85] However, newer emerging data suggest that very high folate intake may actually lead to an increased risk of breast cancer.[86] It is possible that this sort of U-shaped relationship also exists for folate and pregnancy complications. Further investigation is clearly needed.

High Fat or Glucose Intake

It is a well-known fact that Western diets contain disproportionately high amounts of fat and refined sugar. Recent data indicate that a diet high in fat during pregnancy can lead to an increased estrogenic environment for the fetus and therefore cause an increased incidence of mammary gland cancer in the offspring. Normally, as rats age, their mammary estrogen receptor goes through transcriptional silencing via hypermethylation. This may provide protection from cell proliferation and tumor growth. However, when rats are exposed to high-fat diets, their estrogen receptor gene becomes hypomethylated in the promoter region, and this leads to overexpression of the estrogen receptor in rat mammary glands. This seems to be associated with a higher incidence of tumor formation in the exposed offspring.[87]

Additionally, exposure to a high-glucose environment affects the risk of developing type 2 diabetes later in life; therefore, some form of epigenetic memory may persist in insulin/glucose target tissues.[88] It appears that glucose transporter 4, an important transporter in adipose tissue and muscle, is among the genes that are affected. Data show that both the methylation of specific CpG sites and a methylation-sensitive transcription factor play roles in the regulation of the glucose transporter 4 gene during adipocyte differentiation.[88] Also, it appears that there is differential DNA methylation in the promoters of genes implicated in glucose metabolism, such as facilitative glucose transporters and uncoupling proteins.

ENVIRONMENTAL FACTORS

Heavy Metals

Heavy metals are ubiquitous in the environment. It appears that chromium (Cr), cadmium, arsenic, lead, and nickel can promote epigenetic changes in the genome. Cr(III) can modify the epigenome and cause transgenerational carcinogenesis through paternal exposure.[89] When sperm germ cells were exposed to Cr(III), it was found that an allele of the 45S ribosomal RNA spacer promoter was hypomethylated. This may be the mechanism by which carcinogenesis is increased in the offspring.[90] In rat livers, long-term exposure to cadmium increases DNA methylation and DNMT activity and, therefore, leads to tumor formation.[91] In turn, exposure to arsenic can lead to cancer, atherosclerosis, neurological disorders, and cardiovascular disease. In fact, it is classified as a complete transplacental carcinogen in mice and induces cancers of the liver, lungs, bladder, and kidneys. Not only is it a potent mutagen, it also acts through epigenetic mechanisms to alter DNA methylation patterns. Research indicates that arsenic can repress DNMT1 and DNMT3A.[92] Recent data suggest that heavy metals are associated with histone modification and chromatin organization.[93,94]

Endocrine Disruptors

Bisphenol A was first made as an estrogen analogue. Currently, it is widely used in the manufacture of polycarbonate plastics and epoxy resins. It is contained in many types of baby products, such as bottles and pacifiers. It is also present in the plastic sealant used by dentists to prevent cavity formation. Alarmingly, early developmental exposure to bisphenol A can lead to both morphological and functional changes in the male and female genital tracts and mammary glands (thereby causing an increase in susceptibility to infertility) and in the formation of the breast and prostate.[85,86]

2,3,7,8-Tetrachlorodibenzo-*p*-dioxin (TCDD) is a ubiquitous environmental contaminant that is formed during the production of paper, polyvinyl chloride plastics, and chlorinated pesticides and during the incineration of chlorine-containing products. Increased methylation of the IGF2-H19 imprint control region increases ethyltransferase activity, and decreased fetal growth has been shown after the exposure of mouse embryos to TCDD from the 1-cell stage to the blastocyst stage.[95] Furthermore, environmental contaminants such as TCDD,[96] phthalate esters,[97] chlorine disinfectant byproducts,[98] and polychlorinated biphenyls[99,100]can alter DNA methylation, histone modification, and/or steroid hormone metabolism. This also leads to alterations in the reproductive system and the induction of tumor formation.

CONCLUSION

The burgeoning field of epigenetics has provided physicians with new insights into the molecular pathophysiology of the fetal origins of major adult diseases that plague our society, such as obesity, hypertension, diabetes, heart disease, and cancer.

Further research is needed to elucidate how the intrauterine environment of the developing fetus (maternal and external environments) can affect his or her susceptibility to these future diseases, the windows of opportunity for prevention, and the mechanisms controlling the developmental plasticity of epigenetic modifications. Once potentially treatable intrauterine environmental risk factors for human disease are clearly identified, we will have an opportunity to guide the direction and allocation of public health resources and dollars to support the emotional and physical health of pregnant women, their children, and their families for generations to come.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Nafee TM, Farrell WE, Carroll WD, et al. Epigenetic control of fetal gene expression. *BJOG* 2008; 115: 158–168.
2. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. *Science* 2001; 293: 1089–1093.
3. Surani MA. Reprogramming of genome function through epigenetic inheritance. *Nature* 2001; 414: 122–128.
4. Godfrey KM, Lillycrop KA, Burdge GC, et al. Epigenetic mechanisms and the mismatch concept of the developmental origins of health and disease. *Pediatr Res* 2007; 61: 5R–10R.
5. Robertson KD. DNA methylation and human disease. *Nat Rev Genet* 2005; 6: 597–610.
6. Trasler JM. Gamete imprinting: setting epigenetic patterns for the next generation. *Reprod Fertil Dev* 2006; 18: 63–69.
7. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002; 16: 6–21.
8. Quina AS, Buschbeck M, Di Croce L. Chromatin structure and epigenetics. *Biochem Pharmacol* 2006; 72: 1563–1569.
9. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 2006; 31: 89–97.
10. D'Amours D, Desnoyers S, D'Silva I, Poirier GG. Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem J* 1999; 342(pt 2): 249–268.
11. Fuks F, Hurd PJ, Deplus R, Kouzarides T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* 2003; 31: 2305–2312.
12. Geyer PK. The role of insulator elements in defining domains of gene expression. *Curr Opin Genet Dev* 1997; 7: 242–248.
13. Morgan HD, Santos F, Green K, et al. Epigenetic reprogramming in mammals. *Hum Mol Genet* 2005; 14: R47–R58.
14. Beaujean N, Taylor J, Gardner J, et al. Effect of limited DNA methylation reprogramming in the normal sheep embryo on somatic cell nuclear transfer. *Biol Reprod* 2004; 71: 185–193.
15. Beaujean N, Martin C, Debey P, Renard JP. Reprogramming and epigenesis [in French]. *Med Sci (Paris)* 2005; 21: 412–421.
16. Bestor TH. The DNA methyltransferases of mammals. *Hum Mol Genet* 2000; 9: 2395–2402.
17. Richardson B. Impact of aging on DNA methylation. *Ageing Res Rev* 2003; 2: 245–261.
18. Santos F, Hendrich B, Reik W, Dean W. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 2002; 241: 172–182.
19. Dean W, Ferguson-Smith A. Genomic imprinting: mother maintains methylation marks. *Curr Biol* 2001; 11: R527–R530.
20. Ginsburg M, Snow MH, McLaren A. Primordial germ cells in the mouse embryo during gastrulation. *Development* 1990; 110: 521–528.
21. Seki Y, Hayashi K, Itoh K, et al. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol* 2005; 278: 440–458.
22. Sato S, Yoshimizu T, Sato E, Matsui Y. Erasure of methylation imprinting of Igf2r during mouse primordial germ-cell development. *Mol Reprod Dev* 2003; 65: 41–50.
23. Hajkova P, Erhardt S, Lane N, et al. Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* 2002; 117: 15–23.
24. Singh SM, Murphy B, O'Reilly RL. Involvement of gene-diet/drug interaction in DNA methylation and its contribution to complex diseases: from cancer to schizophrenia. *Clin Genet* 2003; 64: 451–460.
25. Waterland RA, Jirtle RL. Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. *Nutrition* 2004; 20: 63–68.
26. Issa JP, Ottaviano YL, Celano P, et al. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. *Nat Genet* 1994; 7: 536–540.
27. McKay JA, Williams EA, Mathers JC. Folate and DNA methylation during in utero development and aging. *Biochem Soc Trans* 2004; 32: 1006–1007.
28. Toyota M, Issa JP. The role of DNA hypermethylation in human neoplasia. *Electrophoresis* 2000; 21: 329–333.
29. Dolinoy DC, Weidman JR, Jirtle RL. Epigenetic gene regulation: linking early developmental environment to adult disease. *Reprod Toxicol* 2007; 23: 297–307.
30. Feil R. Environmental and nutritional effects on the epigenetic regulation of genes. *Mutat Res* 2006; 600: 46–57.
31. Kaati G, Bygren LO, Edvinsson S. Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. *Eur J Hum Genet* 2002; 10: 682–688.
32. Lane N, Dean W, Erhardt S, et al. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* 2003; 35: 88–93.

DOI:10.1002/MSJ

33. Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* 1999; 23: 314–318.
34. Reik W, Walter J. Imprinting mechanisms in mammals. *Curr Opin Genet Dev* 1998; 8: 154–164.
35. Tilghman SM. The sins of the fathers and mothers: genomic imprinting in mammalian development. *Cell* 1999; 96: 185–193.
36. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 2007; 8: 253–262.
37. Badcock C, Crespi B. Imbalanced genomic imprinting in brain development: an evolutionary basis for the aetiology of autism. *J Evol Biol* 2006; 19: 1007–1032.
38. Barker DJ, Osmond C. Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. *Lancet* 1986; 1: 1077–1081.
39. Kensara OA, Wootton SA, Phillips DI, et al. Fetal programming of body composition: relation between birth weight and body composition measured with dual-energy X-ray absorptiometry and anthropometric methods in older Englishmen. *Am J Clin Nutr* 2005; 82: 980–987.
40. Osmond C, Barker DJ, Winter PD, et al. Early growth and death from cardiovascular disease in women. *BMJ* 1993; 307: 1519–1524.
41. Hales CN, Barker DJ. Type 2 (non-insulin-dependent) diabetes mellitus: the thrifty phenotype hypothesis. *Diabetologia* 1992; 35: 595–601.
42. Cooper C, Fall C, Egger P, et al. Growth in infancy and bone mass in later life. *Ann Rheum Dis* 1997; 56: 17–21.
43. Gluckman PD, Hanson MA. The developmental origins of the metabolic syndrome. *Trends Endocrinol Metab* 2004; 15: 183–187.
44. Gluckman PD, Hanson MA. Living with the past: evolution, development, and patterns of disease. *Science* 2004; 305: 1733–1736.
45. Lillycrop KA, Phillips ES, Jackson AA, et al. Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. *J Nutr* 2005; 135: 1382–1386.
46. Lillycrop KA, Slater-Jefferies JL, Hanson MA, et al. Induction of altered epigenetic regulation of the hepatic glucocorticoid receptor in the offspring of rats fed a protein-restricted diet during pregnancy suggests that reduced DNA methyltransferase-1 expression is involved in impaired DNA methylation and changes in histone modifications. *Br J Nutr* 2007; 97: 1064–1073.
47. Burdge GC, Lillycrop KA, Jackson AA, et al. The nature of the growth pattern and of the metabolic response to fasting in the rat are dependent upon the dietary protein and folic acid intakes of their pregnant dams and post-weaning fat consumption. *Br J Nutr* 2008; 99: 540–549.
48. Bogdarina I, Welham S, King PJ, et al. Epigenetic modification of the renin-angiotensin system in the fetal programming of hypertension. *Circ Res* 2007; 100: 520–526.
49. Sinclair KD, Allegrucci C, Singh R, et al. DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. *Proc Natl Acad Sci U S A* 2007; 104: 19351–19356.
50. Weaver IC, Cervoni N, Champagne FA, et al. Epigenetic programming by maternal behavior. *Nat Neurosci* 2004; 7: 847–854.
51. Fraga MF, Ballestar E, Paz MF, et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci U S A* 2005; 102: 10604–10609.
52. Vickers MH, Gluckman PD, Coveny AH, et al. Neonatal leptin treatment reverses developmental programming. *Endocrinology* 2005; 146: 4211–4216.
53. Gluckman PD, Lillycrop KA, Vickers MH, et al. Metabolic plasticity during mammalian development is directionally dependent on early nutritional status. *Proc Natl Acad Sci U S A* 2007; 104: 12796–12800.
54. Wyrwoll CS, Mark PJ, Mori TA, et al. Prevention of programmed hyperleptinemia and hypertension by postnatal dietary omega-3 fatty acids. *Endocrinology* 2006; 147: 599–606.
55. Nylund L, Lunell NO, Lewander R, Sarby B. Uteroplacental blood flow index in intrauterine growth retardation of fetal or maternal origin. *Br J Obstet Gynaecol* 1983; 90: 16–20.
56. Laurin J, Lingman G, Marsal K, Persson PH. Fetal blood flow in pregnancies complicated by intrauterine growth retardation. *Obstet Gynecol* 1987; 69: 895–902.
57. Bell AW, Hay WW Jr, Ehrhardt RA. Placental transport of nutrients and its implications for fetal growth. *J Reprod Fertil Suppl* 1999; 54: 401–410.
58. Waterland RA, Garza C. Potential mechanisms of metabolic imprinting that lead to chronic disease. *Am J Clin Nutr* 1999; 69: 179–197.
59. Maloney CA, Rees WD. Gene-nutrient interactions during fetal development. *Reproduction* 2005; 130: 401–410.
60. Randhawa R, Cohen P. The role of the insulin-like growth factor system in prenatal growth. *Mol Genet Metab* 2005; 86: 84–90.
61. Fowden AL, Sibley C, Reik W, Constancia M. Imprinted genes, placental development and fetal growth. *Horm Res* 2006; 65(suppl 3): 50–58.
62. Constancia M, Hemberger M, Hughes J, et al. Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* 2002; 417: 945–948.
63. Myatt L. Placental adaptive responses and fetal programming. *J Physiol* 2006; 572: 25–30.
64. Roberts JM, Gammill HS. Preeclampsia: recent insights. *Hypertension* 2005; 46: 1243–1249.
65. Granger JP, Alexander BT, Llinas MT, et al. Pathophysiology of preeclampsia: linking placental ischemia/hypoxia with microvascular dysfunction. *Microcirculation* 2002; 9: 147–160.
66. Gupta S, Agarwal A, Sharma RK. The role of placental oxidative stress and lipid peroxidation in preeclampsia. *Obstet Gynecol Surv* 2005; 60: 807–816.
67. Redman CW, Sargent IL. Placental debris, oxidative stress and pre-eclampsia. *Placenta* 2000; 21: 597–602.
68. Elsheikh A, Creatsas G, Mastorakos G, et al. The renin-aldosterone system during normal and hypertensive pregnancy. *Arch Gynecol Obstet* 2001; 264: 182–185.
69. Redman CW, Sargent IL. Preeclampsia and the systemic inflammatory response. *Semin Nephrol* 2004; 24: 565–570.

70. Wang JX, Knottnerus AM, Schuit G, et al. Surgically obtained sperm, and risk of gestational hypertension and pre-eclampsia. *Lancet* 2002; 359: 673–674.

71. Kanayama N, Takahashi K, Matsuura T, et al. Deficiency in p57Kip2 expression induces preeclampsia-like symptoms in mice. *Mol Hum Reprod* 2002; 8: 1129–1135.

72. Knox KS, Baker JC. Genome-wide expression profiling of placentas in the p57Kip2 model of preeclampsia. *Mol Hum Reprod* 2007; 13: 251–263.

73. Chelbi ST, Vaiman D. Genetic and epigenetic factors contribute to the onset of preeclampsia. *Mol Cell Endocrinol* 2008; 282: 120–129.

74. Miller JW, Nadeau MR, Smith J, et al. Folate-deficiency-induced homocysteinaemia in rats: disruption of *S*-adenosylmethionine's co-ordinate regulation of homocysteine metabolism. *Biochem J* 1994; 298(pt 2): 415–419.

75. Xu X, Chen J. One-carbon metabolism and breast cancer: an epidemiological perspective. *J Genet Genomics* 2009; 36: 203–214.

76. Scholl TO, Johnson WG. Folic acid: influence on the outcome of pregnancy. *Am J Clin Nutr* 2000; 71: 1295S–1303S.

77. Shaw GM, Carmichael SL, Nelson V, et al. Occurrence of low birthweight and preterm delivery among California infants before and after compulsory food fortification with folic acid. *Public Health Rep* 2004; 119: 170–173.

78. Knudtson EJ, Smith K, Mercer BM, et al. Serum homocysteine levels after preterm premature rupture of the membranes. *Am J Obstet Gynecol* 2004; 191: 537–541.

79. Wouters MG, Boers GH, Blom HJ, et al. Hyperhomocysteinemia: a risk factor in women with unexplained recurrent early pregnancy loss. *Fertil Steril* 1993; 60: 820–825.

80. Goddijn-Wessel TA, Wouters MG, van de Molen EF, et al. Hyperhomocysteinemia: a risk factor for placental abruption or infarction. *Eur J Obstet Gynecol Reprod Biol* 1996; 66: 23–29.

81. Menon MK, Sengupta M, Ramaswamy N. Accidental haemorrhage and folic acid deficiency. *J Obstet Gynaecol Br Commonw* 1966; 73: 49–52.

82. Dekker GA, de Vries JI, Doelitzsch PM, et al. Underlying disorders associated with severe early-onset preeclampsia. *Am J Obstet Gynecol* 1995; 173: 1042–1048.

83. Laurence KM. Prenatal detection and prevention of neural tube defects in South Wales. *J R Soc Health*. 1986; Oct; 106(5): 153–60.

84. Smithells RW. 1982. Neural tube defects: prevention by vitamin supplements. Pediatrics 69: 498–499.

85. Zhang S, Hunter DJ, Hankinson SE, et al. A prospective study of folate intake and the risk of breast cancer. *JAMA* 1999; 281: 1632–1637.

86. Stolzenberg-Solomon RZ, Chang SC, Leitzmann MF, et al. Folate intake, alcohol use, and postmenopausal breast cancer risk in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *Am J Clin Nutr* 2006; 83: 895–904.

87. Yenbutr P, Hilakivi-Clarke L, Passaniti A. Hypomethylation of an exon I estrogen receptor CpG island in spontaneous and carcinogen-induced mammary tumorigenesis in the rat. *Mech Ageing Dev* 1998; 106: 93–102.

88. Dabelea D, Hanson RL, Lindsay RS, et al. Intrauterine exposure to diabetes conveys risks for type 2 diabetes and obesity: a study of discordant sibships. *Diabetes* 2000; 49: 2208–2211.

89. Cheng RY, Hockman T, Crawford E, et al. Epigenetic and gene expression changes related to transgenerational carcinogenesis. *Mol Carcinog* 2004; 40: 1–11.

90. Shiao YH, Crawford EB, Anderson LM, et al. Allele-specific germ cell epimutation in the spacer promoter of the 45S ribosomal RNA gene after Cr(III) exposure. *Toxicol Appl Pharmacol* 2005; 205: 290–296.

91. Takiguchi M, Achanzar WE, Qu W, et al. Effects of cadmium on DNA-(cytosine-5) methyltransferase activity and DNA methylation status during cadmium-induced cellular transformation. *Exp Cell Res* 2003; 286: 355–365.

92. Reichard JF, Schnekenburger M, Puga A. Long term low-dose arsenic exposure induces loss of DNA methylation. *Biochem Biophys Res Commun* 2007; 352: 188–192.

93. Salnikow K, Costa M. Epigenetic mechanisms of nickel carcinogenesis. *J Environ Pathol Toxicol Oncol* 2000; 19: 307–318.

94. Chen H, Ke Q, Kluz T, et al. Nickel ions increase histone H3 lysine 9 dimethylation and induce transgene silencing. *Mol Cell Biol* 2006; 26: 3728–3737.

95. Edwards TM, Myers JP. Environmental exposures and gene regulation in disease etiology. *Environ Health Perspect* 2007; 115: 1264–1270.

96. Couture LA, Abbott BD, Birnbaum LS. A critical review of the developmental toxicity and teratogenicity of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin: recent advances toward understanding the mechanism. *Teratology* 1990; 42: 619–627.

97. Foster PM. Disruption of reproductive development in male rat offspring following in utero exposure to phthalate esters. *Int J Androl* 2006; 29: 140–147.

98. Tao L, Wang W, Li L, et al. DNA hypomethylation induced by drinking water disinfection by-products in mouse and rat kidney. *Toxicol Sci* 2005; 87: 344–352.

99. Desaulniers D, Xiao GH, Leingartner K, et al. Comparisons of brain, uterus, and liver mRNA expression for cytochrome p450s, DNA methyltransferase-1, and catechol-o-methyltransferase in prepubertal female Sprague-Dawley rats exposed to a mixture of aryl hydrocarbon receptor agonists. *Toxicol Sci* 2005; 86: 175–184.

100. McLachlan JA, Simpson E, Martin M. Endocrine disrupters and female reproductive health. *Best Pract Res Clin Endocrinol Metab* 2006; 20: 63–75.

NOTHING BEATS
BEING FIRST

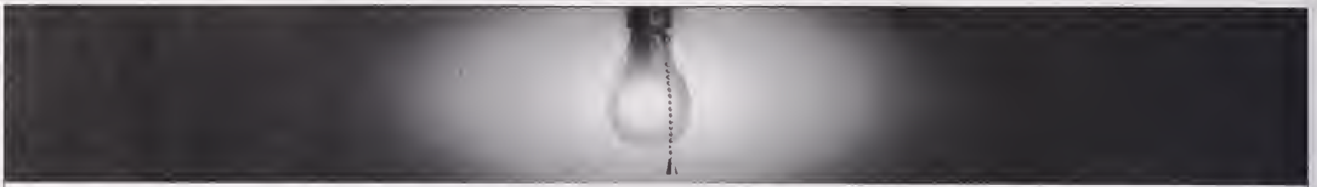
SIGN UP FOR FREE
Journal of Hospital Medicine e-alerts
and be the first to know about new
research in your field!

It's fast, free, and flexible.
Customize your alerts to deliver
only the information you need.

WILEY
InterScience®
DISCOVER SOMETHING GREAT

www.JournalofHospitalMedicine.com

SIGN UP
TODAY!
WILEY-
BLACKWELL
shm
Society of Hospital Medicine



don't publish online in the dark...

let our open linking network shed some light.

This journal is CrossRef-ed! As a scientist or scholar navigating digital content, you've come to expect dynamic linking from citation to original source. As an author, you expect ready dissemination of your work to others. CrossRef is a collaborative, cross-publisher reference linking network that allows users to click on a citation and be taken directly to the full-text location. An independent, non-profit organization that employs open standards, CrossRef was established in cooperation with leading research publishers and the International DOI Foundation. Since the service went live in 2000, CrossRef has linked millions of articles from thousands of journals. Find out more at **www.crossref.org**, and make sure all your electronically published work is CrossRef-enabled.

crossref

CROSSREF.ORG | THE CITATION LINKING BACKBONE

200 WHEELER ROAD, 6TH FLOOR, BURLINGTON, MA 01803 • 781-359-2442

1OUNT SINAI OURNAL OF 1EDICINE

VOL. 77, NO. 3
MAY/JUNE 2010

ournal of Translational and Personalized Medicine

Issue Theme
Endovascular Surgery: Advances in Science and Technique

Theme Editors
Peter Faries, MD
Michael Marin, MD

lominal Aortic Aneurysms:
ovascular Repair

tured Abdominal Aortic
eurysms: Role of
ovascular Therapy

racic Aortic Aneurysms
Dissections:
ovascular Treatment

imally Invasive Vein
gery: Latest Options

acranial Aneurysms:
ovascular Treatment

p Vein Thrombosis:
ovascular Management

ceral Artery Aneurysms:
ovascular Surgery

ecial Features

ard a Person-centered Medicine: From Disease to Patient to Person
rts-related Injuries in Young Female Athletes
sugrel, the Newest Antiplatelet Agent

WILEY-BLACKWELL

Recommend this Journal to Your Library Today!

MOUNT SINAI JOURNAL OF MEDICINE

A Journal of Translational and Personalized Medicine

Editor-in-Chief: **Penny A. Asbell, MD, FACS, MBA**
Mount Sinai School of Medicine, New York, NY

Continuously published since 1934 and included in the *Index Medicus* from the very first volume, the **Mount Sinai Journal of Medicine (MSJM)** has a distinguished career as a general medical journal for a worldwide audience.

Dedicated to improving communication among all participants in medicine today, **MSJM** is an excellent source of up-to-date medical information for healthcare clinicians, educators, researchers and policy makers worldwide. **MSJM** will inspire the creative thinking needed to create new ways to understand, prevent, and treat human disease.

Each issue includes several quality, in-depth review articles related to a single theme and several others on subjects of diverse clinical interest. All theme articles are written by renowned authorities in their fields, from both Mount Sinai and elsewhere, under the supervision of Theme Editors.

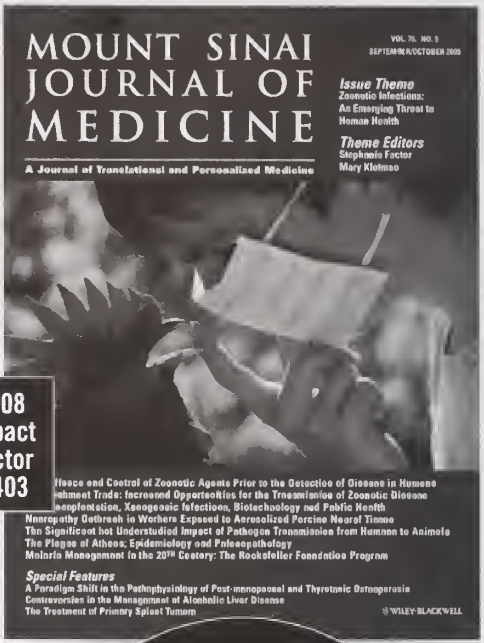
Upcoming Themes for Volume 77:

- Alzheimer's Disease: New Light on Etiology and Management
- Genes, Race, Ethnicity and the Environment
- Endovascular Surgery: Advances in Science and Technique
- Drug Discovery in Academic Settings
- The Obesity Pandemic
- Children's Health and the Environment

ACCESS THESE 2009 EDITORS CHOICE ARTICLES FREE ONLINE

- Advances and Controversies in Assisted Reproductive Technology (76:6)
- Xenotransplantation, Xenogenic Infection, Biotechnology and Public Health (76:5)
- The Student-run Health Clinic: A Novel Arena to Educate Medical Students on Systems-based Practice (76:4)
- Glycemic Targets for Patients with Type 2 Diabetes Mellitus (76:3)
- Neurobehavioral Sequelae of Traumatic Brain Injury (76:2)
- Spectrum of Fibrosing Diffuse Parenchymal Lung Disease (76:1)

Want more? Access a free sample issue online today!



MOUNT SINAI JOURNAL OF MEDICINE

A Journal of Translational and Personalized Medicine

Issue Theme

Theme Editors

Special Features

WILEY-BLACKWELL

2008 Impact Factor 1.403

Stay on Top of the Research!

Sign-up for Email Table of Contents Alerts delivered direct to your desktop – FREE!

Simply go to the homepage and click on "Set E-Mail Alert" for automatic table of contents alerts, or create a customized alert based on your own search criteria.

How to Receive Complimentary Online Access for Institutions?

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Library Recommendation Form

To______________________________

Date______________________________

Requested by:______________________________

Dept.______________________________

I have reviewed the **Mount Sinai Journal of Medicine** and urge you to add a complimentary 2010 online subscription to our subscription list. Please contact me should you have any questions.

WWW.MOUNTSINAIJOURNAL.ORG

WILEY-BLACKWELL

Enter Your 2010 Subscription Today!

MOUNT SINAI JOURNAL OF MEDICINE
Issue Theme
Type 2 Diabetes Mellitus: New Evidence, More Controversy
Theme Editor
Dina Green

med at clinicians and educators, the **ount Sinai Journal of Medicine** provides gh-quality, in-depth review articles in ery issue, studded with clinical pearls.

ur 2010 will include the six 2010 print sues of the **Mount Sinai Journal of edicine** (Volume 77), plus concurrent ll-text online access to these six issues Wiley's Internet publishing platform.

tandard Personal Subscriptions

30.00 (U.S./Canada/Mexico)
20.00 (Outside North America)

ount Sinai Alumni Subscriptions!

0.00 (Worldwide)

xtended for Institutions in 2010

omplimentary nline Access!

ur institution can opt to receive mplimentary online access to the **ount Sinai Journal of Medicine** during 10. When registering for complimentary line access, your institution can also tain one complimentary print bscription.

PERSONAL SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 personal subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $105.00 (U.S./Canada/Mexico) ☐ $130.00 (Outside North America)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

Method of Payment (Canada residents please add GST. Personal rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **John Wiley & Sons, Inc.** Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA

Card #________________ Exp. Date________

Signature________________

(credit card orders not valid without signature)

Ship To:

Name________________

Address________________

City__________ State______ Zip______

MSJM0910

MSJM ALUMNI SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 Alumna/Alumnus Subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $40.00 (Worldwide)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

METHOD OF PAYMENT (Alumna/Alumnus rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **The Mount Sinai School of Medicine.** Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA ☐ Other

Card #________________ Exp. Date________

Signature________________

(credit card orders not valid without signature)

Ship To:

Name________________

Address________________

City__________ State______ Zip______

To subscribe online visit: **www.msjmonline.com**

MSJM10

COMPLIMENTARY ONLINE ACCESS FOR INSTITUTIONS

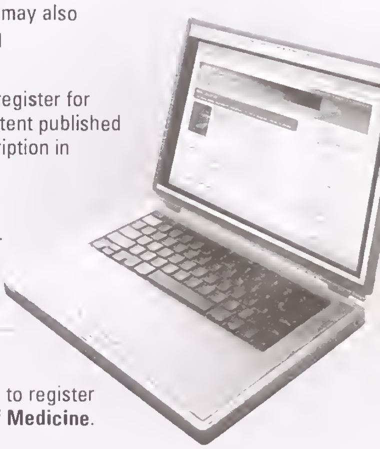
Mount Sinai Journal of Medicine

A Journal of Translational and Personalized Medicine

2010. Volume 77. 6 issues per year. Print ISSN: 0027-2507. Online ISSN: 1931-7581.

Your institution can opt to receive complimentary online access to the **Mount Sinai Journal of Medicine** during 2010. Your institution may also request one complimentary print subscription when registering for complimentary online access.

There is no obligation to subscribe in subsequent years if you register for complimentary access. Institutions retain online access to content published during the complimentary period, with or without a paid subscription in subsequent years.



Dear Librarian,
I have reviewed the scope of Mount Sinai Journal of Medicine.
Please register for complimentary access:

Name: ________ Department: ________

Signature: ________ Date: ________

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Don't delay, recommend to your library today!

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO. 2277 HOBOKEN, NJ

POSTAGE WILL BE PAID BY ADDRESSEE

JOHN WILEY & SONS INC
ATTN: BETH MOLD, MAIL STOP 8-01
111 RIVER STREET
HOBOKEN, NJ 07030-9441

PLACE
POSTAGE
STAMP
HERE

INES SANTOS
MOUNT SINAI JOURNAL OF MEDICINE
ONE GUSTAV L. LEVY PLACE BOX 1094
NEW YORK, NY 10029-6574

Library Routing Card

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

RECEIVED
JUN 15 2010
Levy Library

Editor-in-Chief

Penny A. Asbell

Deputy Editor

David H. Adams

Associate Editors

Stuart Aaronson
Erwin Bottinger
Dennis S. Charney
Valentin Fuster
Wayne Goodman
Jonathan Halperin
Ravi Iyengar
Paul Klotman
Ihor R. Lemischka
Michael L. Marin
Thomas G. McGinn
Ira Nash
Maria New
Rosamond Rhodes
Hugh Sampson
Mary Sano
Jeffrey H. Silverstein
Albert Siu
Anthony Squire
Fred Suchy
Savio Woo

Editorial Board

William J. Bremner	University of Washington
Thomas M. Coffman	Duke University
Dwight Evans	University of Pennsylvania
Gabriel Haddad	University of California, San Diego
Heidi E. Hamm	Vanderbilt University
Matti Häyry	University of Manchester
Eve J. Higginbotham	Howard University
Katherine A. High	University of Pennsylvania
Tony Hope	University of Oxford
Mary Klotman	Duke University
John H. Krystal	Yale University
Raju Kucherlapati	Harvard Medical School
Robert J. Lefkowitz	Duke University
Robert Lisak	Wayne State University
George Lister	University of Texas
Philip A. Marsden	University of Toronto
Charles Nemeroff	Emory University
Scott Rauch	Harvard Medical School
Bernard Roizman	University of Chicago
Alan F. Schatzberg	Stanford University
Inder Verma	The Salk Institute
Jeffrey A. Whitsett	University of Cincinnati
Allan Wolkoff	Albert Einstein College of Medicine
Noam Zohar	Bar-Ilan University

Theme Editors

Peter Faries, MD
Chief, Vascular Surgery
Professor of Surgery and Radiology

Michael Marin, MD
Chair, Department of Surgery
Professor of Surgery and Vascular Surgery

Mount Sinai School of Medicine, New York, NY

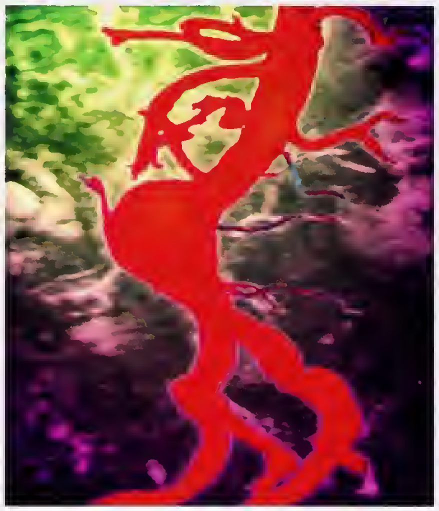
Editorial Director

Stephen Cohen
stephen.cohen@mssm.edu
212-241-6304

Publisher

John Wiley & Sons
Shawn Morton
Journal Publishing Director, Medicine
smorton@wiley.com

Endovascular Surgery: Advances in Science and Technique



This colored magnetic resonance angiography (MRA) scan shows the abdominal arteries (red) of a patient with high blood pressure. At the center is a bulging aortic aneurysm, for which the treatment of choice is endovascular repair. If left untreated, an abdominal aneurysm can rupture, creating an immediate emergency.

Risk factors for aneurysm include smoking, high blood pressure, infection and atherosclerosis. Improved scanning techniques mean that aneurysms are increasingly detected before they cause any symptoms. MRA is a type of magnetic resonance imaging (MRI) which highlights blood vessels without the need to inject a tracer or dye into the patient.

Credit: Zephyr / Photo Researchers, Inc

Future *MSJM* Themes and Theme Editors

- Obesity (Robert Yanagiswa, Derek LeRoith)
- Drug discovery in academic settings (Ravi Iyengar)
- Translational approaches to cancer treatment (William Oh, Steven Burakoff)
- Parkinson's disease (Michele Tagliati, Ron Alterman)
- Children's environmental health (Philip Landrigan)
- Biomarkers in clinical practice (Penny Asbell, Mark Badyatsky)

©2010 **Mount Sinai School of Medicine.** All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, Phone: (978) 750-8400; Fax: (978) 750-4770. Such requests and other permission inquiries should be addressed to the Permissions Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; Phone: (201) 748-6011; Fax: (201) 748-6008; E-mail: permreq@wiley.com. Printed in the United States of America by The Sheridan Press.

Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine (Print ISSN: 0027-2507; Online ISSN: 1931-7581 at Wiley InterScience, www.interscience.wiley.com) is published twelve times per year for the Mount Sinai School of Medicine by Wiley Subscription Services, a Wiley Company, 111 River Street, Hoboken, NJ 07030.

Send subscription inquiries to: Subscription Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-6645.

Offprint sales and inquiries should be addressed to the Reprint Billing Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-8776. All other inquiries should be addressed to the Customer Service Department at (201) 748-6645.

Individual subscription price: $95.00 in North America, $120.00 outside North America. Payment must be made in US dollars drawn on a US bank. Claims for undelivered copies will be accepted only after the following issue has been delivered. For more subscription prices and information, please visit www.mountsinaijournal.org.

Change of address: Please forward six weeks prior to move: present mailing label with change of address. Please enclose a copy of the mailing label or cite your subscriber reference number in order to expedite handling. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits. Send claims in care of: *Mount Sinai Journal of Medicine*, c/o John Wiley & Sons, Inc., Attn: Journals Admin Dept. UK, 111 River Street, Hoboken, NJ 07030.

Cancellations: Subscription cancellations will not be accepted after the first issue has been mailed. Periodicals postage paid at Hoboken, NJ and at additional mailing offices.

Postmaster: Send address changes to *Mount Sinai Journal of Medicine*, Subscription Distribution, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Printed in the United States of America. Single issues from current and recent volumes are available at the current single issue price from cs-journals@wiley.com. Earlier issues may be obtained from Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. Tel: +1 518 537 4700, Fax: +1 518 537 5899, Email: psc@periodicals.com

Advertising inquiries should be addressed to: Advertising Sales Manager, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Phone: (201) 748-6921; Fax: (201) 748-6207. Advertising in this journal is published with a caveat emptor understanding. The Publisher, Editors, and members of the Editorial Board do not imply warranty of products, nor quality, validity, or approval of education material offered by such advertising.

Indexed by: Biological Abstracts® (Thomson ISI), Chemical Abstracts Service/SciFinder (ACS), Current Contents®/Clinical Medicine (Thomson ISI), EMBASE/Excerpta Medica (Elsevier), Index Medicus/MEDLINE/PubMed (NLM), Science Citation Index Expanded™ (Thomson ISI), SCOPUS (Elsevier), Web of Science® (Thomson ISI).

This journal is printed on acid-free paper.

Disclaimer: All articles published, including but not limited to original research, reviews, clinical notes, clinical pearls, prismatic cases, clinician's corner, editorials, reports, and letters, represent the opinions and views of the authors and do not reflect any official policy or medical opinion of the Mount Sinai School of Medicine, the institutions with which the authors are affiliated, or the Publisher unless this is clearly specified. Articles published herein should not be relied upon as recommending or promoting a specific method, diagnosis, or treatment by physicians for any particular patient or disorder. Although the Editors and Publisher believe that drug selections and dosages and the specifications and usage of equipment and devices as set forth herein are in accord with current recommendations and practice at the time of publication, they accept no legal responsibility for any errors or omissions, and make no warranty, express or implied, with respect to material contained herein.

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

CONTENTS
VOL 77, NO 3

Theme Introduction

237 Endovascular Surgery: Advances in Science and Technique
Peter L. Faries, Michael L. Marin

Theme Articles

238 Abdominal Aortic Aneurysms: Endovascular Repair
Jeffrey Jim, Luis A. Sanchez

250 Ruptured Abdominal Aortic Aneurysms: Role of Endovascular Therapy
Neal S. Cayne, Frank J. Veith

256 Thoracic Aortic Aneurysms and Dissections: Endovascular Treatment
Donald T. Baril, Jae S. Cho, Rabih A. Chaer, Michel S. Makaroun

270 Minimally Invasive Vein Surgery: Latest Options for Vein Disease
Steven Elias

279 Intracranial Aneurysms: Endovascular Treatment
Henry Moyle, Aman B. Patel

286 Deep Vein Thrombosis: Endovascular Management
Robert A Lookstein, Christopher F. Giordano

296 Visceral Artery Aneurysms: Review of Current Management Options
Ulka Sachdev-Ost

Special Features

304 Toward Person-centered Medicine: From Disease to Patient to Person
Juan Mezzich, Jon Snaedal, Chris van Weel, Iona Heath

307 Sports-related Injuries in the Young Female Athlete
Alexis C. Colvin, Abigail Lynn

315 Prasugrel, a New Platelet Inhibitor
Debbie Yee, Joanne Meyer, Gina Caliendo, Veronica Oviedo

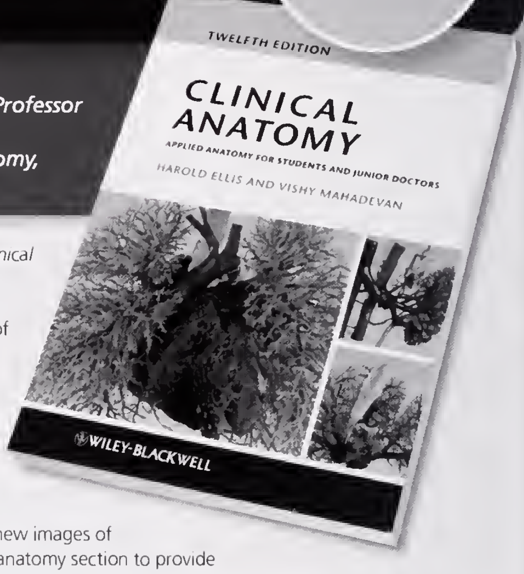
Clinical Anatomy

Applied anatomy for students and junior doctors

TWELFTH EDITION

Harold Ellis and Vishy Mahadevan

Guy's, King's and St Thomas' School of Biomedical Sciences; Emeritus Professor of Surgery, Charing Cross and Westminster Medical School, London; Barbers' Company Reader in Anatomy and Professor of Surgical Anatomy, The Royal College of Surgeons of England, London


TWELFTH EDITION
CLINICAL ANATOMY
APPLIED ANATOMY FOR STUDENTS AND JUNIOR DOCTORS
HAROLD ELLIS AND VISHY MAHADEVAN
WILEY-BLACKWELL

Written by one of the great teachers of Anatomy, this 50th anniversary edition of *Clinical Anatomy* continues to provide thousands of medical students, postgraduate trainees and junior doctors across the world with essential anatomical information within a clinical setting. It is particularly appropriate for those preparing for the Membership of the Royal College of Surgeon's Exam.

For this new edition, Professor Harold Ellis is joined, for the first time, by a co-author. Professor Vishy Mahadevan is Barber's Company Reader in Anatomy at the Royal College of Surgeons, a renowned teacher of surgical trainees and is also an examiner for the MRCS Exam. Together they provide a detailed, easy-to-follow structured text suitable for anatomy students and trainees of all levels.

Fully revised and updated, this new edition features new and re-drawn illustrations, new images of normal anatomy (radiographs, CTs and NMRs) to aid understanding, a revised neuroanatomy section to provide greater clarity and an expansion of the lower limb section.

In order to provide revision support, *Clinical Anatomy* has its own website featuring digital flashcards of all the images found in the textbook available to download for free at **www.ellisclinicalanatomy.co.uk**

March 2010 • 9781405186179
470 pages

CONTENTS

Preface to the Twelfth (Jubilee) Edition
Preface to the First Edition
Acknowledgements to the Twelfth (Jubilee) Edition
Acknowledgements to the First Edition

Part 1: The Thorax
Introduction
Surface anatomy and surface markings
The thoracic cage
The lower respiratory tract
The mediastinum
On the examination of a chest radiograph

Part 2: The Abdomen and Pelvis
Surface anatomy and surface markings
The fasciae and muscles of the abdominal wall
Peritoneal cavity
The gastrointestinal tract
The gastrointestinal adnexae: liver, gall-bladder and its ducts, pancreas and spleen
The urinary tract
The male genital organs
The bony and ligamentous pelvis
The muscles of the pelvic floor and perineum
The female genital organs
The posterior abdominal wall
Computed axial tomography

Part 3: The Upper Limb
Surface anatomy and surface markings of the upper limb
The bones and joints of the upper limb
Three important zones of the upper limb: the axilla, the cubital fossa and the carpal tunnel
The arteries of the upper limb
The brachial plexus
The course and distribution of the principal nerves of the upper limb
Compartments of the upper limb
The female breast
The anatomy of upper limb deformities
The spaces of the hand

Part 4: The Lower Limb
The surface anatomy and surface markings of the lower limb
The bones and joints of the lower limb
Three important zones of the lower limb: the femoral triangle, adductor canal and popliteal fossa
The arteries of the lower limb
The veins of the lower limb
The course and distribution of the principal nerves of the lower limb
Compartments of the lower limb

Part 5: The Head and Neck
The surface anatomy of the neck
The thyroid gland
The palate
The tongue and floor of the mouth
The pharynx
The larynx
The salivary glands
The major arteries of the head and neck
The veins of the head and neck
The lymph nodes of the neck
The cervical sympathetic trunk
The branchial system and its derivatives
The surface anatomy and surface markings of the head
The scalp
The skull
The paranasal sinuses (accessory nasal sinuses)
The mandible
The vertebral column

Part 6: The Nervous System
The brain
The spinal cord
The cranial nerves
The special senses
The autonomic nervous system

Glossary of eponyms
Index

For more information or to place an order visit:
www.ellisclinicalanatomy.co.uk

WILEY-BLACKWELL

COMING SOON!
Download the FREE digital flashcards from *Clinical Anatomy*, at **www.ellisclinicalanatomy.co.uk**

THEME INTRODUCTION

Endovascular Surgery: Advances in Science and Technique

Peter L. Faries, MD, and Michael L. Marin, MD

Department of Surgery, Division of Vascular Surgery, Mount Sinai School of Medicine,
New York, NY

This issue of the *Mount Sinai Journal of Medicine* is dedicated to reviewing the current status of minimally invasive treatments for vascular disease. Tremendous progress has been made in all fields of vascular treatment over the past 2 decades. The result has been that, in many instances, traditional open surgical treatments have been replaced by endovascular approaches. The progress has occurred in parallel across all the areas of vascular disease, ranging from aortic aneurysms and the treatment and prevention of stroke to limb salvage and the treatment of venous disease.

Mount Sinai has been integral to developments and progress in numerous areas of endovascular treatment. Our current chairman of the Department of Surgery, Dr. Michael Marin, is the father of endovascular treatment of aortic aneurysms in the United States. He performed the first endovascular repair of an abdominal aortic aneurysm in North America and went on to develop commercially produced stent grafts dedicated to this technique. Members of our faculty have contributed to the development of percutaneous treatment of carotid disease and intracranial aneurysms for the prevention of stroke. Nationally recognized leaders in the management of coronary artery disease and deep and superficial venous disease and thrombosis who are on the faculty at Mount Sinai have also contributed to this issue.

In addition, pioneers in endovascular treatments from other academic institutions have participated in the preparation of this review. The result reflects the dynamic process that has led to such tremendous improvement in patient care and the associated improvement in the lives of our patients.

As theme editors, Dr. Marin and I would like to thank everyone who has done so much to make this issue a success. We would also like to thank the editors and editorial staff of the *Mount Sinai Journal of Medicine* without whose tireless efforts this issue would not have been completed.

DISCLOSURES

Potential conflict of interest: Dr. Marin is a consultant for Medtronic Corp.

Address Correspondence to:

Peter L. Faries, MD
Department of Surgery,
Division of Vascular Surgery
Mount Sinai School of Medicine
New York, NY
Email: peter.faries@mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20188

© 2010 Mount Sinai School of Medicine

Mount Sinai Journal of Medicine 77:238–249, 2010

Abdominal Aortic Aneurysms: Endovascular Repair

Jeffrey Jim, MD and Luis A. Sanchez, MD

Section of Vascular Surgery, Washington University School of Medicine, St. Louis, MO

OUTLINE

Operative Technique
Endografts
Complications
Clinical Results
  Open Repair Versus Surveillance
  Endovasular Aortic Aneurysm Repair Versus Open Repair
  Endovascular Aortic Aneurysm Repair versus Surveillance
  Ruptured Aneurysms
Society for Vascular Surgery Guidelines for Endovascular Aortic Aneurysm Repair
Future Directions
Conclusion

ABSTRACT

Since its introduction 2 decades ago, endovascular repair of abdominal aortic aneurysms has continued to evolve, and many now consider it the optimal treatment modality for infrarenal abdominal aortic aneurysms. There are currently 5 commercially approved endografts, each with its unique design, and all can be utilized with excellent technical results. The endovascular repair is performed with radiographic guidance utilizing general, spinal, or local anesthesia. Endoleak is the most common complication and unique to the procedure. Routine postoperative surveillance imaging is essential to ensure long-term success. Clinical trials have demonstrated lower perioperative mortality and morbidity for endovascular repair of abdominal aortic aneurysms compared with open repair. However, evidence for long-term survival benefits is lacking. Furthermore, the benefits of endovascular repair in "high-risk" patients remain unclear. Though there is a subset of patients who benefit from endovascular repair of abdominal aortic aneurysms, objective identification of this population remains a challenge. Practitioners currently recommend elective repair for aneurysms $\geq$5.5 cm in diameter and in selected patients with smaller aneurysms. Current studies are underway to evaluate the utility of endovascular repair in small aneurysms. Patients with ruptured aneurysms, due to the high morbidity and mortality associated with their management, may significantly benefit from endovascular repair. Although there is a lack of sufficient randomized controlled trial data, available evidence does support the use of endovascular repair for ruptured abdominal aortic aneurysms. Multiple newer endovascular devices are under clinical investigation to address the limitations of available devices. Ultimately, we may be able to treat the vast majority of patients with all types of abdominal and thoracoabdominal aortic aneurysms with endovascular devices. ***Mt Sinai J Med 77:238–249, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: aneurysm, aorta, endovascular, infrarenal, rupture, surgery.

Address Correspondence to:

Luis A. Sanchez, MD
Section of Vascular Surgery
Washington University School
of Medicine
St. Louis, MO
Email: sanchezl@wudosis.wustl.edu

Abdominal aortic aneurysm (AAA) is a common disease process in developed countries. Risk factors for AAA development include advanced age, tobacco use, male gender, high blood pressure, atherosclerosis, and family history. Operative repair has been the only means of treatment that can prevent rupture and its associated complications. Since the introduction of the techniques described by Dubost in 1951, open surgical repair of AAAs had been considered the gold standard of treatment.[1] Despite numerous advances in critical-care medicine, this operative approach is still associated with significant morbidity and mortality today. Although high-volume tertiary referral centers continue to report excellent

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20180

© 2010 Mount Sinai School of Medicine

outcomes, in-hospital mortality rates of 3% to 4% and complication rates of 30% to 40% can be seen in a sample of hospitals throughout the United States.[2]

The less-invasive techniques of endovascular aortic aneurysm repair (EVAR) were first reported by Parodi and Volodos in 1991.[3,4] Over the next 8 years, these early techniques were used to treat selected high-risk patients with a variety of homemade devices and clinical investigations were performed for the first-generation commercial devices.[5,6] The first 2 commercially available endografts (Ancure device, Guidant, Menlo Park, CA; and AneuRx device, Medtronic, Santa Rosa, CA) in the United States were approved in 1999. For the past decade, the technology has continued to evolve, and there are now 5 US Food and Drug Administration (FDA)–approved endografts available for use, with numerous others under clinical investigation (Table 1). EVAR was quickly adopted and is now widely accepted as an effective, and possibly preferred, way of treating aneurysms located in the infrarenal portion of the aorta. For the past several years, there have been a higher number of aneurysms treated with EVAR than with traditional open repair.[2] In addition, EVAR has allowed physicians to treat a wider population of patients to include those who may have been deemed to have prohibitive risk for open surgical repair.

> *Endovascular aortic aneurysm repair (EVAR) is now widely accepted as an effective, and possibly preferred, way of treating aneurysms located in the infrarenal portion of the aorta.*

OPERATIVE TECHNIQUE

The eligibility of patients for EVAR is dependent on their aortic anatomy (Table 2). Preoperative imaging and planning are essential for technical success and long-term durability of EVAR. With improved noninvasive imaging techniques, most centers now utilize contrast-enhanced spiral computed tomography (CT) as the favored imaging modality. The more invasive contrast angiography is reserved for select situations where interventions are anticipated (e.g., renal artery angioplasty or internal iliac artery embolization) or additional data is needed for complex case planning (e.g., fenestrated endograft use). Several anatomic variables need to be carefully investigated to determine the suitability of endograft use. These include diameter and length of the aortic neck, diameter and length of the iliac arteries, degree of angulation, amount of calcification, thrombus burden, and presence of adequate-sized access vessels. To achieve optimal seal of the endograft to the arterial wall, the infrarenal aortic neck and iliac arteries must be smaller than the expanded endograft diameter by 10% to 15%. An adequate neck length of 10 to 15 mm is also required to prevent migration and subsequent failure. In addition, severe angulation of the infrarenal aortic neck (>60°) or presence of large amounts of calcification or mural thrombus have been shown to be associated with poorer long-term outcomes.

EVAR is performed in the operating room or angiographic suite under fluoroscopic guidance. The patient may receive general, spinal, or local

Table 2. *Optimal Anatomic Characteristics for EVAR.*

Anatomic Feature	Optimal Characteristic
Infrarenal aortic neck	Length ≥15 mm, diameter ≤32 mm, angle ≤60°, minimal thrombus/calcification
Iliac arteries	Fixation site ≥10 mm, diameter ≤20 mm, minimal tortuosity
Access (iliofemoral) vessels	Adequately sized to accommodate introducer system

Abbreviations: EVAR, endovascular aortic aneurysm repair.

Table 1. *Endografts in the United States.*

Withdrawn	FDA-Approved (Current)	Under Investigation
Ancure*	AneuRx (Medtronic, Santa Rosa, CA)	Anaconda (Vascutek, Renfrewshire, Scotland)
Vanguard	Excluder (W.L. Gore, Flagstaff, AZ)	Aorfix (Lombard, Oxfordshire, England)
Lifepath	Powerlink (Endologix, Irvine, CA)	Aptus (Aptus Endosystems, Sunnyvale, CA)
Fortron	Talent (Medtronic, Santa Rosa, CA)	Endurant (Medtronic)
Enovus	Zenith (Cook Medical, Bloomington, IN)	UniFit (LeMaitre Vascular, Burlington, MA)
		Zenith LP (Cook Medical)
		Zenith Fenestrated (Cook Medical)

Abbreviations: FDA, US Food and Drug Administration; LP, low-profile.
* Initial FDA approval was granted in 1999, and it was voluntarily withdrawn by the manufacturer in 2003.

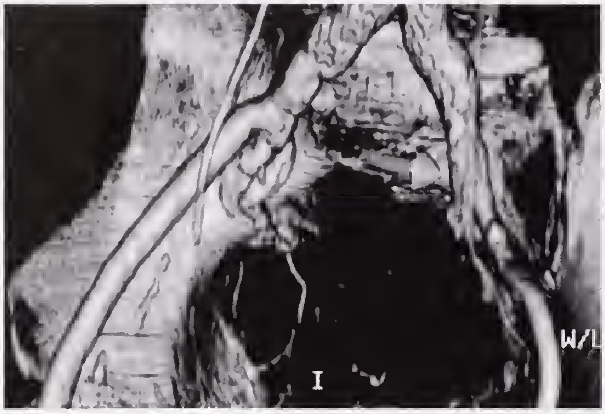


Fig 1. A postoperative computed tomography scan demonstrating a right-sided iliac conduit. After its use during endovascular aortic aneurysm repair for the delivery of the endograft, the conduit was utilized as a common iliac artery to common femoral artery bypass to address the patient's concomitant external iliac artery occlusion.

anesthesia. Standard arterial access for endograft delivery is via surgical exposure of the bilateral common femoral arteries. However, a purely percutaneous approach is advocated by some and can be utilized with excellent results in selected patients.[7] For patients with inadequately sized femoral vessels, a surgical conduit (typically a 10 mm Dacron graft) may be sewn directly to an iliac vessel of suitable size (exposed via a retroperitoneal dissection) to aid in the delivery of the endograft to the aorta (Figure 1). Aortic anatomy is confirmed with the use of contrast angiography, carbon-dioxide angiography, or intravascular ultrasound. The surgical goal is to deliver the endograft in the proper location to exclude blood flow from the aneurysm sac. The proximal covered end of the main body endograft is deployed proximal to the aneurysm just inferior to the orifice of the renal arteries. The distal limbs of the stent graft are guided into the iliac arteries and deployed separately with generous component overlap in the modular systems. The device is then molded and secured in position using a molding balloon. Completion angiography is utilized to confirm proper placement of the endograft and to rule out major endoleaks (Figure 2). Though hospital length of stay differs for each patient, most may be discharged home within 24 hours after the procedure.

Though hospital length of stay differs for each patient, most may be discharged home within 24 hours after the procedure.

ENDOGRAFTS

There are currently 5 FDA-approved endografts commercially available in the United States. Each device has its unique design and method of endograft fixation to the aortic wall (Table 3). Consequently, each

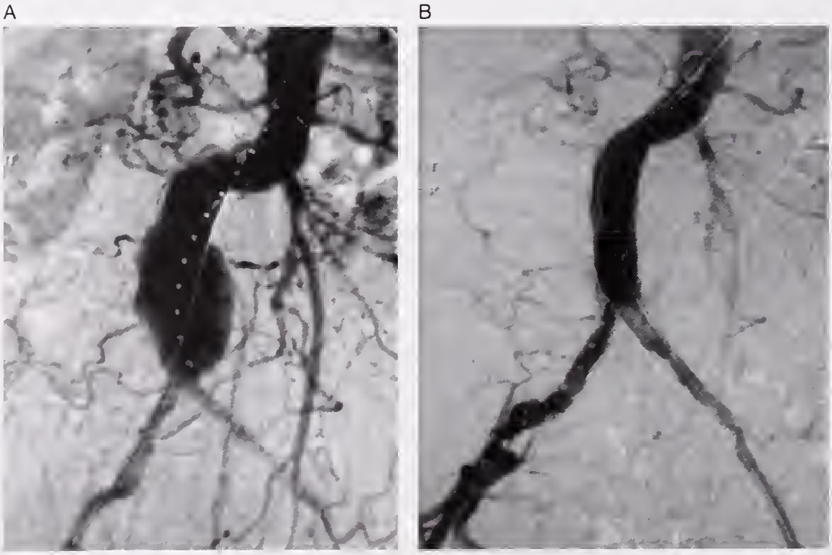


Fig 2. (A) Intraoperative angiography demonstrating infrarenal abdominal aortic aneurysm (AAA). (B) Completion angiography showing successful deployment of endograft, patent renal and mesenteric vessels, and no evidence of major endoleak.

Table 3. *Device-Specific Graft Design.*

Endograft	Components	Materials	Method of Fixation
AneuRx	Modular (contralateral limb)	Nitinol, polyester	Radial force, column strength
Excluder	Modular (contralateral limb)	Nitinol, ePTFE	Radial force, main body barbs
Zenith	Modular (2 docking limbs)	Stainless steel, Dacron	Radial force, suprarenal stent with barbs
Powerlink	Unibody bifurcated	Cobalt-chromium alloy, ePTFE	Radial force, "anatomic fixation" at bifurcation
Talent	Modular (contralateral limb)	Nitinol, polyester	Radial force, bare suprarenal stent

Abbreviations: ePTFE, expanded polytetrafluoroethylene.

Table 4. *Device-Specific Long-Term Outcomes of EVAR at 5 Years.*

Endograft	Freedom from ARM	Freedom from Conversion	Freedom from Rupture
AneuRx	97%	90%	98%
Excluder	97%	96%	100%
Zenith	98%*, 96%†	99%*, 98%†	100%*, 100%†
Powerlink	98%	‡	100%
Talent	97%	99%	98%

NOTE: Data are from studies in references 8 to 12.
Abbreviations: ARM, aneurysm-related mortality; EVAR, endovascular aortic aneurysm repair.
* Standard-risk group.
† High-risk group.
‡ The 6-year follow-up data revealed only 1 case of conversion in the entire trial.

endograft has its own specific set of FDA-approved instructions for use, which may differ from those of the other available devices. With appropriate patient selection and accurate graft deployment, excellent technical results can be achieved with low perioperative mortality rates of <2% and high rates of long-term freedom from rupture, aneurysm-related mortality, and conversion (Table 4).[8–12]

> *With appropriate patient selection and accurate graft deployment, excellent technical results can be achieved with low perioperative mortality rates of <2% and high rates of long-term freedom from rupture, aneurysm-related mortality, and conversion.*

COMPLICATIONS

Endoleaks are problems uniquely related to EVAR. The term "endoleak" was introduced to describe an incomplete exclusion of blood flow from the aneurysm sac.[13] Persistent blood flow in the aneurysm sac outside of the endograft is the most frequent complication after EVAR and has been reported in 1 in 4 patients during follow-up.[14] However, most of them resolve on their own during the early follow-up period. Persistent endoleaks can lead to aneurysm growth with an increased risk of subsequent rupture. The development of endoleaks is one of the primary reasons why postprocedure surveillance is essential for long-term clinical success of this procedure. Endoleaks can be classified into 4 different categories (Table 5). Type I and III endoleaks mandate surgical intervention, as they are associated with a higher risk of aneurysmal rupture. These are due to failure of the proximal or distal seal zones of the device (type Ia and type Ib) or failure

Table 5. *Types of Endoleaks.*

Classification	Cause(s)
Type I	Inadequate seal of proximal (Ia) or distal (Ib) extent of endograft to aortic wall Inadequate seal from iliac occluder (Ic)
Type II	Back perfusion from patent lumbar, middle sacral, inferior mesenteric, hypogastric, or accessory renal artery
Type III	Modular disconnection (IIIa) or fabric tear/disruption (IIIb)
Type IV	Fabric porosity

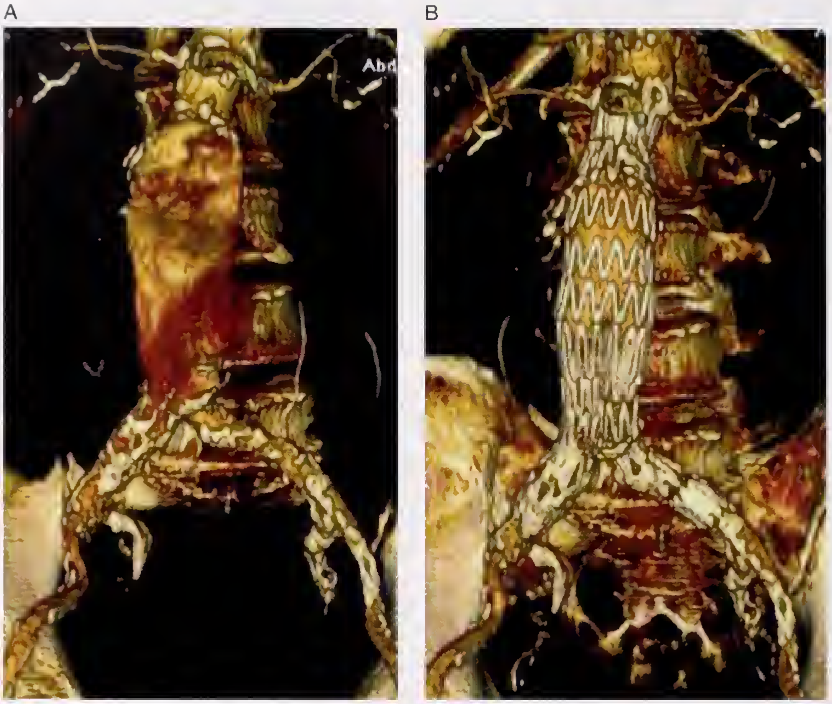


Fig 3. (A) Preoperative computed tomography (CT) scan demonstrating intact infrarenal abdominal aortic aneurysm. (B) Postoperative CT scan at 6 months showing proper placement of endograft with aneurysm exclusion and no evidence of endoleak.

at the junction of modular components (type III). Type II endoleaks, due to inflow into the aneurysm sac from patent inferior mesenteric or lumbar arteries, often are benign findings that resolve during follow-up. Intervention for type II endoleaks is only warranted in cases of increase in aneurysm size.[15] Type IV endoleaks due to porosity of the graft material are rare with the newer generations of endografts.

Whereas EVAR compares favorably with open surgical repair, systemic complications remain a concern with EVAR. These range from cardiac and pulmonary dysfunction to renal failure, embolic disease, and spinal cord ischemic injury. In addition to endoleaks, EVAR also presents other unique access-related and endograft-related complications not encountered in open surgical repair. Early complications can include iatrogenic arterial injuries, unintentional coverage of critical branch vessels (e.g., renal artery), and endograft limb thrombosis. Late complications can occur, and thus use of EVAR mandates continued radiographic endograft surveillance. These can include endoleaks, graft migration, limb thrombosis, and implant-related complications such as fractures, component separation, and fabric tears. Long-term surveillance is essential to ensure durable clinical results. Repeat imaging with contrast-enhanced CT angiography is typically performed at 1 month, 6 months and annually post device implantation (Figure 3). A reduced surveillance regimen may be appropriate in instances where there is early success with newer devices. This may improve patient safety by reducing the effects of intravenous contrast and radiation exposure, while also reducing health care costs. However, this has not been validated in a randomized prospective-trial setting. In patients where contrast use is prohibited (e.g., renal insufficiency, allergy), then duplex ultrasound may be combined with noncontrast CT for complete evaluation. Magnetic resonance angiography may also be utilized in select patients. Surveillance imaging is critical to evaluate for the development of endoleaks, limb kinking/stenosis, limb occlusion, and migration. This allows treatment of these problems, often in an endovascular fashion, to maintain aneurysm exclusion.

CLINICAL RESULTS

Open Repair Versus Surveillance

The single most important risk factor for AAA rupture is maximum aneurysm diameter. The risk of rupture is substantial for patients with aneurysms of at least 5.5 cm in diameter and increases exponentially with increasing diameter.[16] Two large randomized controlled trials investigated whether prophylactic open surgical repair for small aneurysms was beneficial. The United Kingdom Small Aneurysm Trial (UKSAT) and the subsequent Aneurysm Detection and Management (ADAM) Trial both evaluated patients with intact small aneurysms of 4.0 cm to 5.5 cm in size.[17–20] Both studies randomized patients into 2 groups: (1) immediate/early repair, and (2) selective/delayed repair with initial surveillance, followed by repair in cases of growth to 5.5 cm, rapid enlargement, or development of symptoms.

The UKSAT enrolled 1090 patients aged 60 to 76 with no medical contraindication to open surgical repair.[17,18] The 30-day operative mortality for the early-surgery group was 5.8%, which led to a survival disadvantage for these patients early in the trial. Approximately three-quarters (73.8%) of the surveillance group eventually had aneurysm repair by 8 years of follow-up. The survival curves of the 2 groups crossed at 3 years, and by 8 years the mortality rate in the early-surgery group was 7.2% lower ($P = 0.03$) than in the surveillance group. The authors concluded that among patients with small AAAs, there was no long-term difference in mean survival between the 2 groups. Although the mortality rate was lower in the early-surgery group, the difference was felt to be partially attributed to beneficial lifestyle changes (smoking cessation) by members of the early-surgery group. A major criticism of this study was the high perioperative mortality rate of 5.8%. Many felt that if surgical repair could be performed with a lower incidence of perioperative mortality, follow-up of the early-surgery group would show a distinct survival benefit. The ADAM trial analyzed 1136 patients aged 50 to 79 years with AAA of 4.0 to 5.4 cm who did not have high surgical risk for immediate open repair.[19,20] The mean follow-up was 4.9 years, with a range of 3.5 to 8.0 years. Even with a lower (compared with the UKSAT) total operative mortality of 2.7% in the immediate-repair group, the trend in survival did not favor immediate repair in any subgroups analyzed. This study also had a high crossover rate, as a large proportion (61.6%) of patients in the surveillance group ultimately required surgical repair. Whereas both studies demonstrated that immediate/early surgery for patients with intact small aneurysms did not confer a long-term survival benefit, it did show that most patients will ultimately require surgical intervention. In addition, the study did not separate patients with aneurysms <5 cm from those 5 to 5.5 cm, which would better reflect the generally accepted treatment criteria for intervention on patients with aneurysms >5 cm. Nevertheless, most practitioners thus supported a policy of reserving elective repair for AAAs at least 5.5 cm in diameter as long as there are no other anatomic or medical factors (e.g., female gender, poorly controlled hypertension, family history, current smoking history, transplant patients, and COPD) that suggest an increased risk of aneurysm rupture at a smaller size.

Nevertheless, most practitioners have supported a policy of reserving elective repair for abdominal aortic aneurysms (AAAs) at least 5.5 cm in diameter as long as there are no other anatomic or medical factors (e.g., female gender, poorly controlled hypertension, family history, current smoking history, transplant patients, and COPD) that suggest an increased risk of aneurysm rupture at a smaller size.

Endovasular Aortic Aneurysm Repair Versus Open Repair

When EVAR was introduced, it was viewed as an alternative for patients who were considered high risk for open surgical repair. The early clinical experience not only demonstrated technical feasibility of EVAR, but the patients also had shorter hospital stays with lower perioperative complication and mortality rates. As EVAR gained wider clinical application, investigators began to pose several important clinical questions: is EVAR comparable to open surgical repair in all patients? For the subgroup of high-risk patients, did they receive any benefit from undergoing EVAR? Furthermore, with EVAR being performed with a lower perioperative mortality rate than open repair (UKSAT and ADAM trials), is there survival benefit for early intervention and should the threshold for repair decrease from 5.5 cm?

"Medically Fit" Population

The Dutch Randomized Endovascular Aneurysm Management (DREAM) Trial is a multicenter study that compared open repair with endovascular repair in 351 patients with AAAs of at least 5 cm who were considered suitable candidates for both techniques.[21,22] The operative mortality was 4.6% for the open-repair group and 1.2% in the endovascular-repair group (risk ratio 3.9). The combined rate of operative mortality and severe complications was 9.8% in the open-repair group and 4.7% in the endovascular-repair group (risk ratio 2.1). These initial results led the investigators to report that endovascular repair was preferable to open repair in patients with AAAs of at least 5 cm in size.[21] However, further follow-up indicated that the perioperative survival advantage with endovascular repair was limited to the first postoperative year and not sustained over time. By 2 years of follow-up, both groups had a 66% rate of survival free of moderate or severe complications. The EVAR-1 trial randomized 1082 medically fit elective patients aged at least 60 years with aneurysms ≥5.5 cm diameter.[23,24] The 30-day mortality rate in the EVAR group was 1.7%, versus 4.7% for open repair ($P < 0.01$). At 4 years after randomization, despite a higher proportion of postoperative complications in the EVAR group (41% versus 9%, $P < 0.0001$), there was a persistent reduction in aneurysm-related deaths in the EVAR group (4% versus 7%, $P = 0.04$). However, this came at a higher mean hospital cost, with negligible difference in health-related quality of life (HRQL). Furthermore, EVAR also offers no advantage with long-term all-cause mortality rates (26% versus 29%, $P = 0.46$). The authors thus concluded that EVAR offers no advantage with respect to all-cause mortality and HRQL, is more expensive, and leads to a greater number of complications and reinterventions. However, it does result in a 3% improvement in aneurysm-related survival.

As devices and techniques continue to evolve and improve, and in light of the relatively high perioperative mortality rates in the European trials, the Open Versus Endovascular Repair (OVER) Veterans Affairs Cooperative Study Group began a study in 2002 to evaluate EVAR in the United States.[25] The study randomized 881 veterans aged ≥49 years with minimum aneurysm diameter of 5.0 cm. The perioperative mortality was lower for endovascular repair compared with open repair (0.5% versus 3.0%, $P = 0.004$). At 2 years, there was no significant difference in mortality (7.0% versus 9.8%, $P = 0.13$), major morbidity, secondary therapeutic procedures, and HRQL. In discussing their results, the authors attributed the lower perioperative mortality rates for both treatments compared with the European trials to several factors. The OVER trial procedures were performed from 2002 to 2007, compared with 1999 to 2003 for the DREAM and EVAR-1 trials. Furthermore, differences in surgical technique and postoperative care cannot be excluded, nor the differences in endovascular systems used dismissed. Nevertheless, even though open repair was performed with low mortality, endovascular repair resulted in fewer perioperative deaths, and this early advantage was not offset by increased morbidity and mortality in the first 2 years after repair.

The combination of all these results has led to the increased application of EVAR to patients with AAAs who have suitable anatomy for endovascular repair. In addition, many physicians have broadened the application of this technology to patients with borderline anatomy for endovascular repair in view of the periprocedural advantages. The long-term results for this latter group of patients are still unknown and may or may not be comparable with the excellent results obtained with approved devices in patients with suitable aortic anatomy. Currently, in patients with suitable anatomy and who are medically fit, the options of endovascular and open surgical repair, with their advantages and disadvantages, are discussed and considered. The vast majority of patients select endovascular treatment due to its early perioperative advantages and the less-invasive nature of the procedure.

Currently, in patients with suitable anatomy and who are medically fit, the options of endovascular and open surgical repair, with their advantages and disadvantages, are discussed and considered. The vast majority of patients select endovascular treatment due to its early perioperative advantages and the less-invasive nature of the procedure.

High-Risk Population

The EVAR-2 trial was initiated to identify whether EVAR improved survival compared with no intervention in patients unfit for open AAA repair.[26] The study randomized 338 patients aged 60 years or older who had aneurysms of at least 5.5 cm in diameter. All had poor health status and were considered unfit for major surgery. Patients were assigned either EVAR (n = 166) or no intervention (n = 172) and all received best medical therapy. The findings showed that the patients had significantly worse health than those enrolled in EVAR-1. This resulted in a high 30-day mortality rate of 9% in the EVAR group and an overall mortality after 4 years of 64%. There was no significant difference between the 2 groups in all-cause or aneurysm-related mortality. The mean hospital costs per patient over the 4 years of follow-up were significantly higher in the EVAR group with no difference in the HRQL score. The results led the authors to conclude that EVAR did not improve survival over observation and was associated with a need for continued surveillance and reinterventions, at a substantially increased cost. Since its publication, several important criticisms have been made regarding the results of EVAR-2 and thus its clinical applicability. The aneurysm-related mortality of 14% in the EVAR group was markedly higher than that reported in other high-risk series. In the study, 14 patients died in the interval between randomization and repair. Thus, 52% of the perioperative deaths as well as 19% of the total deaths in the EVAR group occurred preoperatively. Nine patients died from aneurysm rupture before their surgery date, and this accounts for 45% of the 20 aneurysm-related deaths in the EVAR group. The median time from randomization to rupture for these 9 patients was 98 days (range 6–767). In a group of patients with a mean aneurysm diameter of 6.4 cm, a median wait period of 57 days for the EVAR group does not coincide with our standard medical practice and may have falsely inflated the death rate in the treatment arm. Furthermore, more than a quarter of patients (47 of 172) assigned to no intervention subsequently underwent aneurysm repair. At least 18 of these patients were intervened upon for rupture, accelerated growth, or development of symptoms. This clearly reduces the aneurysm-related death in the no-intervention group.

In an effort to evaluate the outcome of EVAR and open surgery for high-risk patients, there are 2 large studies comparing these 2 treatment modalities in the United States. The Society for Vascular Surgery (SVS) Outcomes Committee analyzed data collected from 5 multicenter investigational device exemption clinical trials.[27] Using high-risk criteria as defined by EVAR-2, 565 EVAR patients and 61 open surgical controls were identified. The 30-day operative mortality was 2.9% in EVAR and 5.1% in the open group ($P = 0.32$). The aneurysm-related death rate after EVAR was 3.0% at 1 year and 4.2% at 4 years, compared with 5.1% at both time points for the open group ($P = 0.58$). The overall survival at 4 years after EVAR was 56%, versus 66% in the open group ($P = 0.23$). After treatment, EVAR successfully prevented rupture in 99.5% at 1 year and 97.2% at 4 years. The authors thus concluded that EVAR in anatomically suited high-surgical-risk patients is safe and provides lasting protection from aneurysm-related mortality. An important limitation of this study was the low number of patients in the open group, and the study was admittedly underpowered to identify a difference in mortality between the 2 modalities. In a study using prospectively collected information from the Department of Veterans Affairs National Surgical Quality Improvement Program (NSQIP), a cohort of high-risk patients was evaluated to compare EVAR (788 patients) and open repair (1580 patients).[28] Elective EVAR had a lower 30-day (3.4% versus 5.2%, $P = 0.047$) and 1-year (9.5% versus 12.4%, $P = 0.038$) all-cause mortality than open repair. The risk of perioperative complications was significantly lower after EVAR (16.2% versus 31.0%, $P < 0.0001$). The retrospective analysis led the authors to argue that patients with considerable medical comorbidities and AAA benefit from and should be considered for primary EVAR. These 2 studies differ from EVAR-2, as there was no surveillance arm in either study. Most practitioners will agree that there are many high-risk patients who benefit from endovascular repair. However, there is a subset of high-risk patients who will not benefit from either EVAR or open repair. In the absence of further prospective randomized trials, the challenge remains objective identification of this subgroup.

Endovascular Aortic Aneurysm Repair versus Surveillance

As discussed above, surgical intervention for AAA ≥ 5.5 cm is widely accepted, though it is common practice to start considering treatment of patients with AAAs ≥ 5.0 cm. However, randomized controlled trials have failed to demonstrate a survival benefit for early intervention with open surgical repair on aneurysms of a lesser size. However, as the outcomes of EVAR continue to improve with low risk of perioperative mortality, some question if the threshold for intervention should be lowered. Two device-specific prospective randomized clinical trials are underway to evaluate early EVAR versus surveillance for smaller AAAs. The European-based multicenter Comparison of Surveillance Versus Aortic Endografting for Small Aneurysm Repair (CAESAR) Trial will randomize patients to surveillance or early EVAR with the Zenith Device.[29] The study started in September 2004 with a target recruitment of 740 patients with aneurysms of 4.1 to 5.4 cm diameter. The primary endpoint for the study is all-cause mortality at 54 months. However, due to slow recruitment, the trial closed with only 360 patients randomized and follow-up is continuing at this point. The 70-site PIVOTAL (Positive Impact of Endovascular Options for Treating Aneurysm Early) trial in the United States completed enrollment in

November 2008. The study randomized 728 patients with small AAAs (4–5 cm) for surveillance or EVAR with the AeurRx or Talent stent-grafts.[30] The interim results with a mean follow-up of 20 months of this study have been presented.[31] There was no difference in rupture, aneurysm-related morality, or survival between the 2 groups. The PIVOTAL trial ended patient recruitment early, as an interim review of the data suggested that complete accrual of the planned number of patients would not generate results that were any different from the results with the patients already enrolled. Until further results of these 2 trials are available, the optimal diameter for early treatment with EVAR remains unknown. Unfortunately, none of these studies evaluated with enough power the more clinical relevant subgroup of patients in the range of 5.0 to 5.4 cm that may be considered for treatment currently.

Ruptured Aneurysms

Most ruptured AAAs (rAAAs) are treated with conventional open repair. However, despite advances in critical-care medicine, open repair is still associated with a high mortality rate (almost 50%).[32] The first EVAR for rAAA was reported in 1994.[33] Despite a lack of long-term randomized controlled trial (RCT) results to properly evaluate EVAR compared with open repair for rAAAs, its use has slowly increased over time (from 5.9% in 2001 to 18.9% in 2006) with varying degrees of success.[34] The available outcome data are heterogeneous and have a variety of inherent biases. In most studies, the use of EVAR often is limited to those patients with hemodynamic stability and in centers with personnel and technology equipped for emergent EVAR. Furthermore, it is generally estimated that less than half of rAAAs have anatomic characteristics favorable for use with conventional endografts.[35] There is certainly a significant component of publication bias also. The completion of RCTs to resolve these issues has proved difficult and likely to be impossible in the United States. The Nottingham randomized trial was abandoned after it was determined to be underpowered to demonstrate any significant difference.[36] The Amsterdam Acute Aneurysm Trial began in 2004 with planned recruitment until 2007. After enrolling 80 patients, there was no statistically significant difference in the primary combined endpoint of 30-day mortality and severe morbidity. The recruitment has thus been extended twice.[37] The Paris-based Ruptured Aorta-Iliac Aneurysms: Endo versus Surgery (ECAC) trial is designed to recruit only patients who are hemodynamically stable and morphologically suitable for EVAR.[38] The randomization process is unusual and based on calendar week and hospital at which the patient presents. The United Kingdom–based Immediate Management of the Patient with Rupture: Open Versus Endovascular Repair (IMPROVE) Trial began enrollment recently and plans for randomization of 600 patients.[39] The study will address whether a strategy of preferential EVAR versus open repair will significantly reduce the 30-day and in-hospital mortality in rAAA.

Until completion of the RCTs noted above, published results are limited to single-center experiences, population-based/database analysis, and systemic reviews of the literature.[40] However, the available data do suggest the utility of EVAR for rAAAs. Analysis of 1044 pairs of matched patients from the Medicare population showed a survival benefit of EVAR that persisted over 4 years of follow-up.[41] The survival rates were also improved with increased surgeon and hospital volume. In evaluating 27,750 hospital discharges for rAAA, EVAR had a lower overall in-hospital mortality than open repair (31.7% versus 40.7%, $P < 0.0001$).[34] Higher elective open-repair volume as well as rAAA volume increased this mortality advantage further. In an effort to collect a large sample of the world experience with EVAR for rAAA, data from 49 centers committed to EVAR treatment whenever possible were collected for analysis.[42] The study gathered information on 1037 patients treated by EVAR and 763 patient treated by open repair. The overall 30-day mortality in all EVAR patients was 21.2%. In centers performing EVAR whenever possible, almost half (49%) of the patients underwent EVAR. The 30-day mortality rate for these centers was 19.7% for EVAR (680 patients) compared with 36.3% for open repair (763 patients; $P < 0.0001$).

Randomized controlled trials of EVAR for ruptured AAAs are in process. While the results are pending, an effort was made to collect a large sample of the world experience. Data were collected for analysis from 49 centers committed to EVAR treatment whenever possible. The 30-day mortality rate for these centers was 19.7% for EVAR (680 patients) compared with 36.3% for open repair (763 patients; $P < 0.0001$).

These and other similar findings have led some to advocate a move toward regionalization and centers of excellence to decrease the mortality rate of repaired rAAAs. Furthermore, establishing a protocol for urgent and emergent EVAR for rAAAs appears to be essential to obtain optimal results.[43] Finally, it is worth mentioning that abdominal compartment syndrome (ACS) is a major cause of morbidity and mortality after EVAR for rAAA. Use of aortic occlusion balloon, coagulopathy, massive transfusion requirements, and conversion from bifurcated to aortouniiliac devices are all factors associated with the development of ACS.[44] Early recognition through measurements of bladder pressure and surgical decompression are necessary to further improve survival.

SOCIETY FOR VASCULAR SURGERY GUIDELINES FOR ENDOVASCULAR AORTIC ANEURYSM REPAIR

The SVS recently published practice guidelines for the care of patients with AAAs.[45] The document provides important recommendations regarding the general approach to a patient with an AAA, treatment options, management in the perioperative and postoperative periods, as well as highlighting areas in need of further research. In commenting on the indications for EVAR, there was a strong recommendation that emergent EVAR should be considered for treatment of a ruptured AAA if anatomically feasible. For patients unfit for open repair, the ability of EVAR to provide a survival advantage remains uncertain and the overall quality of evidence is low. To ensure continued excellent outcomes in light of widespread adoption and utility of EVAR, the SVS strongly recommends that EVAR should be performed at centers with documented in-hospital mortality for elective EVAR of <3% and a perioperative conversation rate to open repair of <2%.

To ensure continued excellent outcomes in light of widespread adoption and utility of EVAR, the Society for Vascular Surgery strongly recommends that EVAR should be performed at centers with documented in-hospital mortality for elective EVAR of <3% and a perioperative conversation rate to open repair of <2%.

FUTURE DIRECTIONS

EVAR can be performed with excellent technical results in patients with appropriate anatomy using currently available devices. However, for those with unfavorable proximal aortic neck characteristics, use of available endografts outside the manufacturer's instruction for use results in lower technical success with a higher rate of endograft-related adverse events (e.g., migration, endoleak).[46] Multiple newer endovascular devices are currently under clinical investigation to try to expand the applicability of EVAR and address the limitations of the available devices (Table 1). Lower-profile systems are being developed with the hope of reducing the need for surgical conduits or vascular access–related complications. The concern for endograft migration and proximal attachment failure is being addressed by new-generation devices (Endurant, Anaconda) with the addition of barbs or hooks for active fixation. The Aptus endograft utilizes a unique endostapling apparatus for secure endograft fixation. The UniFit endovascular graft utilizes an aortouniiliac design with bare suprarenal stents. Severe aortic neck angulation is another characteristic often associated with graft migration or poor endograft apposition to the aortic wall with resultant type Ia endoleaks. The Aorfix and Anaconda endografts have been developed with circular supporting rings that add flexibility to the device while maintaining structural support. The Aorfix device is currently under investigation to evaluate its performance in AAAs with a severely angulated neck of >60°. Short aortic necks are also associated with an increased risk of migration and type Ia endoleak. Endovascular treatment of these patients requires active fixation within the normal suprarenal portion of the aorta with endograft seal across and above the renal arteries. This can be accomplished with custom-made fenestrated endografts that are specifically designed to accommodate the perivisceral anatomy of patients with short aortic necks. These devices have been used worldwide and a recent report of the intermediate-term (24 months) results from a multicenter study in the United States demonstrated reasonable outcomes.[47]

CONCLUSION

EVAR continues to evolve and is becoming the treatment of choice for patients with suitable anatomy based on the excellent published long-term results. The technology continues to rapidly improve to address the limitation of currently approved devices. In some selected experienced vascular centers, up to

80% of patients with infrarenal AAAs are already being treated with approved and investigational devices. The recently published SVS practice guidelines for EVAR serve as an excellent summary of the current status of this technology. Ultimately, we may be able to expand the use of EVAR to treat the vast majority of patients with all types of abdominal and thoracoabdominal aortic aneurysms with endovascular devices.

DISCLOSURES

Potential conflict of interest: Dr. Sanchez has consulting relationships with the following companies: Cook Medical, Medtronic Inc, W.L. Gore, Aptus endosystems, Trivascular II.

REFERENCES

1. Dubost C, Allary M, Oeconomos N. Resection of an aneurysm of the abdominal aorta: reestablishment of the continuity by a preserved human arterial graft, with result after five months. *AMA Arch Surg* 1952; 64: 405–408.
2. Schwarze ML, Shen Y, Hemmerich J, Dale W. Age-related trends in utilization and outcome of open and endovascular repair for abdominal aortic aneurysm in the United States, 2001–2006. *J Vasc Surg* 2009; 50: 722–729.
3. Parodi JC, Palmaz JC, Barone HD. Transfemoral intraluminal graft implantation for abdominal aortic aneurysm. *Ann Vasc Surg* 1991; 5: 491–499.
4. Volodos NL, Karpovich IP, Troyan VI, et al. Clinical experience of the use of self-fixing synthetic prostheses for remote endoprosthetics of the thoracic and the abdominal aorta and iliac arteries through the femoral artery and as intraoperative endoprosthesis for aorta reconstruction. *Vasa Suppl* 1991; 33: 93–95.
5. Marin ML, Veith FJ, Cynamon J, et al. Initial experience with transluminally placed endovascular grafts for the treatment of complex vascular lesions. *Ann Surg* 1995; 222: 449–465.
6. Moore WS, Rutherford RB; for EVT Investigators. Transfemoral endovascular repair of abdominal aortic aneurysm: results of the North American EVT phase 1 trial. *J Vasc Surg* 1996; 23: 543–553.
7. Lee WA, Brown MP, Nelson PR, Huber TS. Total percutaneous access for endovascular aortic aneurysm repair ("Preclose" technique). *J Vasc Surg* 2007; 45: 1095–1101.
8. Zarins CK; for AneuRx Clinical Investigators. The US AneuRx Clinical Trial: 6-year clinical update 2002. *J Vasc Surg* 2003; 37: 904–908.
9. Peterson BG, Matsumura JS, Brewster DC, Makaroun MS; for Excluder Bifurcated Endoprosthesis Investigators. Five-year report of a multicenter controlled clinical trial of open versus endovascular treatment of abdominal aortic aneurysms. *J Vasc Surg* 2007; 45: 885–890.
10. Greenberg RK, Chuter TA, Cambria RP, et al. Zenith abdominal aortic aneurysm endovascular graft. *J Vasc Surg* 2008; 48: 1–9.
11. Wang GJ, Carpenter JP; for Endologix Investigators. The Powerlink system for endovascular abdominal aortic aneurysm repair: six-year results. *J Vasc Surg* 2008; 48: 535–545.
12. Sanchez LA; for Investigators of the Talent eLPS IDE Trial. Five-year outcomes of the Medtronic Talent AAA Stent-Graft Pivotal IDE Trial. *J Vasc Surg* 2009; 50: 148.
13. White GH, Yu W, May J. Endoleak–a proposed new terminology to describe incomplete aneurysm exclusion by an endoluminal graft. *J Endovasc Surg* 1998; 5: 378–380.
14. Ouriel K, Clair DG, Greenberg RK, et al. Endovascular repair of abdominal aortic aneurysms: device-specific outcome. *J Vasc Surg* 2003; 37: 991–998.
15. Van Marrewijk C, Buth J, Harris PL, et al. Significance of endoleaks after endovascular repair of abdominal aortic aneurysms: the EUROSTAR experience. *J Vasc Surg* 2002; 35: 461–473.
16. Lederle FA, Johnson GR, Wilson SE, et al. Rupture rate of large abdominal aortic aneurysms in patients refusing or unfit for elective repair. *JAMA* 2002; 287: 2968–2972.
17. UK Small Aneurysm Trial Participants. Mortality results for randomised controlled trial of early elective surgery or ultrasonographic surveillance for small abdominal aortic aneurysms. *Lancet* 1998; 352: 1649–1655.
18. UK Small Aneurysm Trial Participants. Long-term outcomes of immediate repair compared with surveillance of small abdominal aortic aneurysms. *N Engl J Med* 2002; 346: 1445–1452.
19. Powell JT, Brown LC, Forbes JF, et al. Final 12-year follow-up of surgery versus surveillance in the UK Small Aneurysm Trial. *Br J Surg* 2007; 94: 702–708.
20. Lederle FA, Wilson SE, Johnson GR, et al. Immediate repair compared with surveillance of small abdominal aortic aneurysms. *N Engl J Med* 2002; 346: 1437–1444.
21. Prinssen M, Verhoeven EL, Buth J, et al. A randomized trial comparing conventional and endovascular repair of abdominal aortic aneurysms. *N Engl J Med* 2004; 351: 1677–1679.
22. Blankensteijn JD, de Jong SECA, Prinssen M, et al. Two-year outcomes after conventional or endovascular repair of abdominal aortic aneurysms. *N Engl J Med* 2005; 352: 2398–2405.
23. Greenhalgh RM, Brown LC, Kwong GP; for EVAR Trial Participants. Comparison of endovascular aneurysm repair with open repair in patients with abdominal aortic aneurysm (EVAR trial 1), 30-day operative mortality results: randomised controlled trial. *Lancet* 2004; 364: 843–848.
24. EVAR Trial Participants. Endovascular aneurysm repair versus open repair in patients with abdominal aortic aneurysm (EVAR trial 1): randomised controlled trial. *Lancet* 2005; 365: 2179–2186.
25. Lederle FA, Freischlag JA, Kyriakides TC, et al.; for Open Versus Endovascular Repair (OVER) Veterans Affairs Cooperative Study Group. Outcomes following endovascular vs open repair of abdominal aortic aneurysm: a randomized trial. *JAMA* 2009; 302: 1535–1542.
26. EVAR Trial Participants. Endovascular aneurysm repair and outcome in patients unfit for open repair of abdominal aortic aneurysm (EVAR trial 2): randomised controlled trial. *Lancet* 2005; 365: 2187–2192.

27. Sicard GA, Zwolak RM, Sidawy AN, et al. Endovascular abdominal aortic aneurysm repair: long-term outcome measures in patients at high risk for open surgery. *J Vasc Surg* 2006; 44: 229–236.
28. Bush RL, Johnson ML, Hedayati N, et al. Performance of endovascular aortic aneurysm repair in high-risk patients: results from the Veterans Affairs National Surgical Quality Improvement Program. *J Vasc Surg* 2007; 45: 227–234.
29. CAESAR Trial Collaborators. Comparison of surveillance versus Aortic Endografting for Small Aneurysm Repair (CAESAR) trial: study design and progress. *Eur J Vasc Endovasc Surg* 2005; 30: 245–251.
30. Ouriel K. The PIVOTAL study: a randomized comparison of endovascular repair versus surveillance in patients with smaller abdominal aortic aneurysms. *J Vasc Surg* 2009; 49: 266–269.
31. Ouriel K, Clair DG, Kent KC, Zarins CK; for PIVOTAL Investigators. Endovascular repair compared with surveillance for patients with small abdominal aortic aneurysms. *J Vasc Surg* 2009; 50: 449.
32. Brown MJ, Sutton AJ, Bell PR, Sayers RD. A meta-analysis of 50 years of ruptured abdominal aortic aneurysm repair. *Br J Surg* 2002; 89: 714–730.
33. Yusuf SW, Whitaker SC, Chuter TA, et al. Emergency endovascular repair of leaking aortic aneurysm. *Lancet* 1994; 344: 1645.
34. McPhee J, Eslami MH, Arous EJ, et al. Endovascular treatment of ruptured abdominal aortic aneurysms in the United States (2001–2006): a significant survival benefit over open repair is independently associated with increased institutional volume. *J Vasc Surg* 2009; 49: 817–826.
35. Slater BJ, Harris EJ, Lee JT. Anatomic suitability of ruptured abdominal aortic aneurysms for endovascular repair. *Ann Vasc Surg* 2008; 22: 716–722.
36. Hinchliffe RJ, Bruijstens L, MacSweeney ST, Braithwaite BD. A randomised trial of endovascular and open surgery for ruptured abdominal aortic aneurysm–results of a pilot study and lessons learned for future studies. *Eur J Vasc Endovasc Surg* 2006; 32: 506–513.
37. Amsterdam Acute Aneurysm Trial Collaborators. Amsterdam Acute Aneurysm trial: background, design, and methods. *Vascular* 2006; 14: 130–135.
38. National Institutes of Health. Ruptured Aorta-Iliac Aneurysms: Endo vs. Surgery (ECAR) Trial. Available at: http://www.clinicaltrials.gov; ID: NCT00577616. Accessed 8 November 2009.
39. Powell JT. Time to IMPROVE the management of ruptured abdominal aortic aneurysm: IMPROVE trialists. *Eur J Vasc Endovasc Surg* 2009; 38: 237–238.
40. Hinchliffe RJ, Powell JT, Cheshire NJ, Thompson MM. Endovascular repair of ruptured abdominal aortic aneurysm: a strategy in need of definitive evidence. *J Vasc Surg* 2009; 49: 1077–1080.
41. Egorova N, Giacovelli J, Greco G, et al. National outcomes for the treatment of ruptured abdominal aortic aneurysm: comparison of open versus endovascular repairs. *J Vasc Surg* 2008; 48: 1092–1100.
42. Veith FJ, Lachat M, Mayer D, et al. Collected world and single center experience with endovascular treatment of ruptured abdominal aortic aneurysms. *Ann Surg* 2009; 250: 818–824.
43. Mehta M, Taggert J, Darling RC III, et al. Establishing a protocol for endovascular treatment of ruptured abdominal aortic aneurysms: outcomes of a prospective analysis. *J Vasc Surg* 2006; 44: 1–8.
44. Mehta M, Darling RC III, Roddy SP, et al. Factors associated with abdominal compartment syndrome complicating endovascular repair of ruptured abdominal aortic aneurysms. *J Vasc Surg* 2005; 42: 1047–1051.
45. Chaikof EL, Brewster DC, Dalman RL, et al. The care of patients with an abdominal aortic aneurysm: the Society for Vascular Surgery practice guidelines. *J Vasc Surg* 2009; 50: 2S–49S.
46. Abbruzzese TA, Kowlek CJ, Brewster DC, et al. Outcomes following endovascular abdominal aortic aneurysm repair (EVAR): an anatomic and device-specific analysis. *J Vasc Surg* 2008; 48: 19–28.
47. Greenberg RK, Sternbergh WC III, Makaroun M, et al. Fenestrated Investigators. Intermediate results of a United States multicenter trial of fenestrated endograft repair for juxtarenal abdominal aortic aneurysms. *J Vasc Surg* 2009; 50: 730–737.

Ruptured Abdominal Aortic Aneurysms: Role of Endovascular Therapy

Neal S. Cayne, MD,[1] and Frank J. Veith, MD[1,2]

[1]New York University Medical Center, New York, NY
[2]Cleveland Clinic, Cleveland, OH

OUTLINE

Techniques for Endovascular Repair of Ruptured Abdominal Aortic Aneurysms
  Hypotensive Hemostasis and Restricted Fluid Resuscitation
  Proximal Balloon Control
  Advantages of Endovascular Treatment of Ruptured Abdominal Aortic Aneurysms
Collected World Experience with Endovascular Graft Treatment of Ruptured Abdominal Aortic Aneurysms
  Why Results of Endovascular Repair for Ruptured Abdominal Aortic Aneurysms Vary
Discussion

ABSTRACT

Ruptured abdominal aortic aneurysms historically have high mortality rates. Despite improvements in many open surgical techniques and perioperative care, these mortality rates have not significantly changed. Some of the reasons for the high mortality rates include the excessive blood loss and hypothermia that occur during open operative repair. The blood loss and hypothermia, combined with resuscitative dilutional coagulopathy, can lead to an irreversible spiraling coagulopathy that ultimately ends in the patient's demise. The availability of endovascular approaches to treat abdominal aortic aneurysms in the early 1990s offered an opportunity to substantially alter the treatment outcomes of ruptured abdominal aortic aneurysms. Endovascular repair offers many advantages, including rapid aortic control under local anesthesia, as well as an opportunity to limit the hypothermia and blood loss that occur with an open abdomen. This article will review the endovascular management of ruptured abdominal aortic aneurysms and describe the endovascular techniques for safe and effective treatment. ***Mt Sinai J Med 77:250–255, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: abdominal aortic aneurysm, aneurysm, balloon control, endograft, hypotensive hemostasis, rupture.

Ruptured abdominal aortic aneurysms (RAAAs) historically have high mortality rates. Approximately 50% of patients with RAAAs do not make it to the hospital alive, and of those who do, 30% to 55% die during or after standard open surgical repair.[1–5] Despite improvements in many open surgical techniques and perioperative care, these mortality rates have not significantly changed.[1–6] Some of the reasons for the high mortality rates include the excessive blood loss and hypothermia that occur during open operative repair. The blood loss and hypothermia, combined with resuscitative dilutional coagulopathy, can lead to an irreversible spiraling coagulopathy that ultimately ends in the patient's demise. During open repair, it is important to obtain rapid aortic control to limit blood loss. This aortic control was thought to most rapidly be obtained by laparotomy with sequential placement of a supracelic and infrarenal aortic cross-clamp.[7] This, unfortunately, requires general anesthesia, which can cause profound hypotension and cardiac arrest if the patient is hypovolemic upon induction.

Address Correspondence to:

Neal S. Cayne
New York University
New York, NY
Email: neal.cayne@nyumc.org

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20182

© 2010 Mount Sinai School of Medicine

> *During open repair of ruptured abdominal aortic aneurysms (RAAAs), it is important to obtain rapid aortic control to limit blood loss. Unfortunately, the preferred procedure– laparotomy with sequential placement of a supracelic and infrarenal aortic cross-clamp– requires general anesthesia, which can cause profound hypotension and cardiac arrest if the patient is hypovolemic upon induction.*

The availability of endovascular approaches to treat abdominal aortic aneurysms (AAAs) in the early 1990s offered an opportunity to substantially alter the treatment outcomes of RAAAs.[8,9] Endovascular repair offers many advantages, including rapid aortic control under local anesthesia, as well as an opportunity to limit the hypothermia and blood loss that occur with an open abdomen. Limiting the blood loss and hypothermia can help avoid the spiraling cascade of coagulopathy, leading to additional blood loss and hypothermia, worsening the coagulopathy. The first reported endovascular treatment for AAA was described in 1991 by Parodi.[10] Since then, the procedure and techniques have been refined. In addition, there are currently numerous commercial grafts that are readily available. This article will review the endovascular management of RAAAs and describe the endovascular techniques for safe and effective treatment.

> *Endovascular repair offers many advantages in the treatment of ruptured abdominal aortic aneurysms (RAAAs), including rapid aortic control under local anesthesia, as well as an opportunity to limit the hypothermia and blood loss that occur with an open abdomen.*

TECHNIQUES FOR ENDOVASCULAR REPAIR OF RUPTURED ABDOMINAL AORTIC ANEURYSMS

Hypotensive Hemostasis and Restricted Fluid Resuscitation

It has been established that in circumstances with major arterial bleeding, restricted blood transfusions and fluid resuscitation can actually decrease blood loss and improve patient outcomes.[11–14] As early as 1991, fluid restriction was advocated for the preoperative management of RAAAs.[15] When patients are over-resuscitated, the pressure elevates and a formed clot at the site of rupture or arterial injury can get dislodged, causing more bleeding. In addition, if excessive fluids are used, coagulation factors get diluted, which can contribute to a coagulopathy. We routinely advocate strict restriction of fluid and blood resuscitation for RAAAs. The use of vasopressors is avoided, and we accept a blood pressure as low as 50 mm Hg if the patient is moving and responsive. As long as this is the case, fluid restriction is continued until the patient is in the operating room. At that point, access for proximal balloon control can be obtained under local anesthesia. If the patient becomes unresponsive with an unobtainable blood pressure, fluid and blood resuscitation may become necessary while obtaining proximal balloon control. Patients with RAAAs frequently deteriorate with the induction of general anesthesia, as the anesthetics cause vasodilatation and hypotension. Therefore, general anesthesia is ideally avoided until aortic control is obtained. Once in the operating room, if the patient is thin, exposure of the femoral arteries can usually be performed under local anesthesia. The procedure can then be commenced via these incisions if the patient remains stable, avoiding general anesthesia until the graft is in place (or possibly at all). If this is not possible, either femoral or brachial access needs to be obtained and endovascular tools for proximal balloon control must be readily available prior to induction.

Proximal Balloon Control

If patients deteriorate, the femoral (or brachial) access is rapidly exchanged for a 14- or 16-Fr sheath (Figure 1). Heparin can then be given. A compliant balloon (30 to 40 mm) can be inserted into the sheath and inflated with dilute contrast under fluoroscopic control in either the supraceliac or pararenal aorta. If the femoral route is used, a 30-cm-long sheath is kept in place to prevent caudal migration of the balloon secondary to aortic pulsations (Figure 2). With

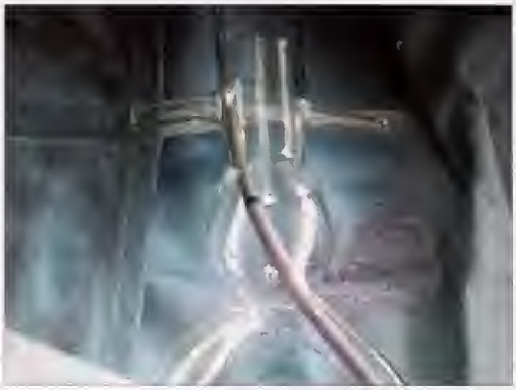


Fig 1. A 16-Fr sheath placed via femoral access.

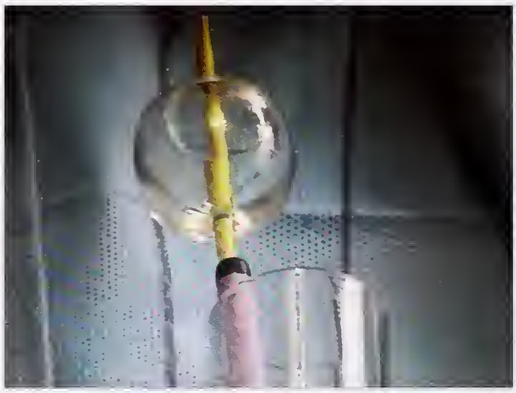


Fig 2. Inflated compliant balloon with 16-Fr sheath to prevent caudal migration.

the balloon inflated, the remainder of the procedure is conducted as rapidly as possible to minimize the degree of visceral and renal ischemia (Figure 3A, 3B). If the infrarenal neck is not suitable for an endovascular repair, balloon control can be maintained until the abdomen is opened and aortic control, preferably below the renal arteries, can be obtained. If a bifurcated endograft is planned, a compliant balloon should be placed in the infrarenal neck inside of the graft as soon as possible, to limit renal and visceral ischemia (Figure 4). The remainder of the procedure is completed in an expeditious manner.

Advantages of Endovascular Treatment of Ruptured Abdominal Aortic Aneurysms

Among the advantages of endovascular repair of ruptured aneurysms are the ability to obtain proximal control without general anesthesia and the ability to deploy the graft from a remote access site. In addition, blood loss is reduced and hypothermia minimized by eliminating laparotomy, thus limiting a spiraling, uncontrolled coagulopathy.

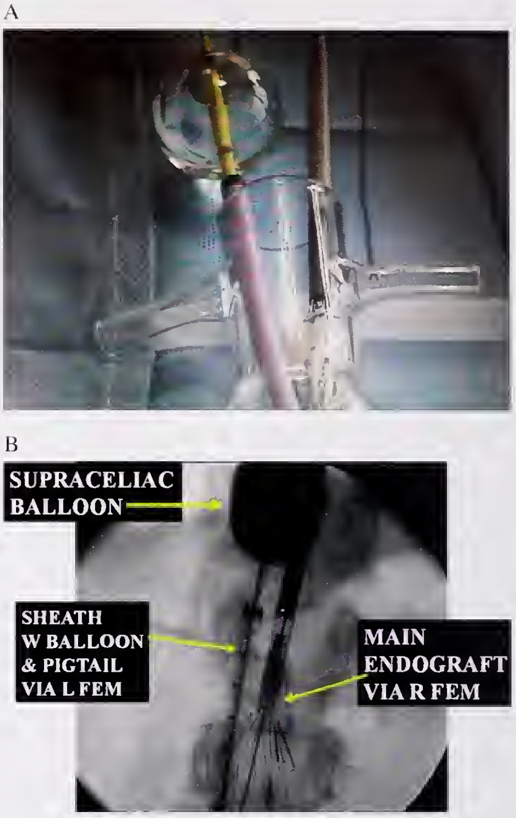


Fig 3. (A) Inflated compliant balloon above partially deployed endovascular device. (B) Fluoroscopy image of inflated compliant balloon with partially deployed endovascular graft. **Abbreviations:** W, with; L FEM, left femoral artery; R FEM, right femoral artery.

Proximal Control without General Anesthesia

Patients with ruptured AAAs may be severely hypotensive. However, many patients may have their blood pressure stabilized at a nonlethal level. This is due to sympathetically mediated vasoconstriction

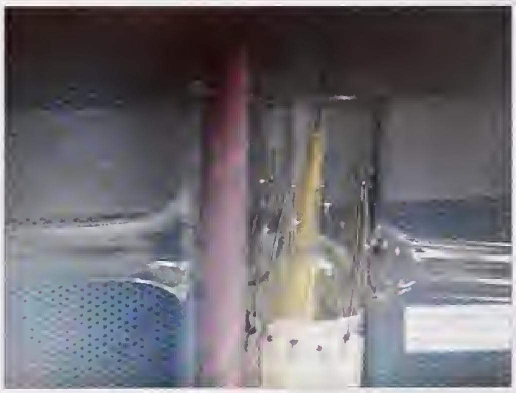


Fig 4. Compliant balloon placed inside of the endovascular graft below the renal arteries to allow renal and visceral perfusion.

in response to hypotension. It is not uncommon for this vasoconstriction to be released during the induction of general anesthesia, which results in a sudden drop in blood pressure. Therefore, a relatively stable patient may become severely hypotensive, mandating urgent application of a proximal aortic clamp. However, a guidewire can be inserted in the upper abdominal or lower thoracic aorta through a percutaneous puncture under local anesthesia, while maintaining the vasoconstriction. Once the guidewire is inserted in the aorta, the patient can then safely undergo induction of general anesthesia because proximal control can be rapidly and relatively safely obtained by placement of a large sheath and an occlusion balloon using the previously inserted guidewire.

Deployment of Graft from a Remote Access Site

Endovascular grafts are inserted and deployed through a remote access site, thereby obviating the need for laparotomy and, more importantly, eliminating the technical difficulties that are encountered when performing a standard repair in the rupture setting. With the associated bleeding, the anatomy of the retroperitoneal structures is often distorted and obscured by a large hematoma, which may lead to technical difficulties as well as inadvertent injury of the inferior vena cava, the left renal vein or its genital branches, the duodenum, or other surrounding structures. These iatrogenic injuries have been the cause of significant operative morbidity and mortality following standard surgery for RAAAs. In contrast, endograft repair is performed within the arterial tree, which is unaffected by extravasated blood or previous operative scarring.

Reduced Blood Loss

Endovascular repair for RAAAs has been accomplished with a relatively small amount of additional blood loss compared with that which occurs during open RAAA repair. This advantage is more important in patients with RAAAs because these patients have already lost a significant amount of blood following rupture, and coagulopathy or disseminated intravascular coagulation secondary to further blood loss can be serious and often lethal complications. There are several reasons why blood loss is limited, including the maintenance of the tamponade effect within the retroperitoneum. In addition, back-bleeding from the iliac and lumbar arteries and bleeding from the anastomotic suture lines and from iatrogenic venous injuries are eliminated.

Minimized Hypothermia

Hypothermia secondary to poor perfusion and laparotomy can exacerbate coagulopathy, which is one of the causes of mortality following open surgical repair. Endovascular graft repair can minimize the extent of hypothermia by avoiding laparotomy.

COLLECTED WORLD EXPERIENCE WITH ENDOVASCULAR GRAFT TREATMENT OF RUPTURED ABDOMINAL AORTIC ANEURYSMS

Over the last 9 years, a collaborative group, the Endovascular Aneurysm Repair (EVAR) Investigators, has been pooling its results with the use of endovascular graft repair of RAAAs. The results of this study were published in the November 2009 issue of *Annals of Surgery*. Many of the details of this experience cannot be described here; however, the highlights can be summarized. Data were collected from 44 centers on 1037 RAAA patients who were treated by EVAR. The overall 30-day mortality in these patients was 21%, although that figure may be biased by the fact that in many centers worse-risk, unstable patients were uniformly treated by open surgery, whereas EVAR was used on more stable patients. However, 13 centers were identified in which EVAR was employed to treat all anatomically suitable RAAA patients. In these 13 centers, 680 RAAA patients treated by EVAR had a 30-day mortality of 19.7%, whereas 763 RAAA patients treated by open repair had a 30-day mortality of 36.3%. Although the 2 patient groups may have not been perfectly comparable, this mortality difference is highly statistically significant ($P < .0001$). In addition,

> *Using data collected from 13 centers over 9 years, a group of investigators compared endovascular aneurysm repair (EVAR) with open repair in the treatment of all anatomically suitable RAAA patients. The data showed that 680 RAAA patients treated by EVAR had a 30-day mortality rate of 19.7%, whereas 763 RAAA patients treated by open repair had a 30-day mortality rate of 36.3% ($P < .0001$).*

many of the EVAR-treated patients were prohibitive risks for a standard open repair. The low mortality for endograft repair coupled with the inclusion of many high-risk RAAA patients strongly suggests that endovascular graft repair, when feasible, will improve treatment outcomes for RAAA.

Why Results of Endovascular Repair for Ruptured Abdominal Aortic Aneurysms Vary

Although some groups have achieved excellent results with EVAR for RAAAs and have reported a low 30-day mortality ranging from 10% to 25%, other centers have not been able to reproduce these results. Indeed, some reports have shown no improvement when comparing EVAR with open repair in the treatment of RAAAs. We believe these variable results are due to technical factors, the use of adjuncts, the strategies used in the performance of the EVAR procedures themselves, and myriad other details, including AAA volume, that may influence the outcome of treatment for RAAAs. Key among these are the use of a protocol and organization for the treatment of RAAAs, the presence of an endovascularly trained staff committed to this protocol, and a readily available stock of endovascular grafts. Also very important in the optimal management of RAAAs are the routine use of hypotensive hemostasis, the awareness of abdominal compartment syndrome and a willingness to detect it early and treat it aggressively,[16] and appropriate use with proper techniques for supraceliac aortic balloon control. All of these components of treatment are important in achieving optimal results with EVAR in the treatment of RAAAs. We believe that when all these components of treatment are in place, EVAR will produce better outcomes with RAAAs than open repair, as has been demonstrated in a number of centers that employ them. If these components of treatment are used, the superiority of EVAR will become apparent, and a randomized prospective trial will not only be unnecessary but will also be impossible to perform for ethical and logistic reasons.

DISCUSSION

The aforementioned techniques for endovascular treatment of RAAAs offer the ability to obtain proximal control without general anesthesia, to deploy the graft from a remote access site, and to minimize blood loss, hypothermia, and coagulopathy.

It is encouraging that patients with RAAAs treated endovascularly have relatively low mortality rates, particularly because many are high-risk patients who are poor surgical candidates.[17–20] These results show that endograft repair of RAAAs is feasible and effective in selected cases. However, before the universal use of these techniques is adopted, many questions must be answered with certainty. Should endovascular repair be attempted in all RAAA patients, or only in those who are relatively stable? In what proportion of RAAA patients should an endograft repair be used? Should a CT scan be obtained before RAAA patients are taken to the operating room or endovascular suite? In which patients should a proximal balloon be placed? How should this be done? Is a randomized prospective comparison with open repair necessary or justified? Should an aortounilateral graft be used or a modular bifurcated graft? What kind of anesthesia should be employed? What resources are required for an institution to undertake endograft repairs of RAAAs? Until these questions are answered, we will not know the optimal approach to endovascular repair of RAAAs. Nevertheless, we believe that endovascular grafts represent a potentially better way to treat this entity in light of the fact that previous open surgical methods have had such a persistently high morbidity and mortality. We currently believe that these endovascular grafts can and should be used in all anatomically suitable patients and in more than 50% to 60% of all patients who present with a RAAA. Indeed, some of the sickest, most urgent hypotensive patients may benefit the most from endograft treatment rather than subjecting them to an emergent open repair under suboptimal conditions. Moreover, we believe that the use of fluoroscopic techniques to facilitate the placement of proximal occlusion balloons for RAAAs,[17,20–22] an old idea,[23–26] will make this endovascular adjunct a practical and valuable one, even if an endovascular graft procedure is not possible and an open repair is required. And finally, we believe that hypotensive hemostasis or restricted fluid resuscitation will prove valuable in the RAAA setting, whether treatment is by EVAR or open repair. Hypotensive hemostasis will likely become the standard of care for all RAAAs, leading to improved treatment outcomes for this lethal condition.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Johansen K, Kohler TR, Nicholls SC, et al. Ruptured abdominal aortic aneurysm: the Harborview experience. *J Vasc Surg* 1991; 13: 240–247.

DOI:10.1002/MSJ

2. Gloviczki P, Pairolero PC, Mucha P. Ruptured abdominal aortic aneurysms: repair should not be denied. *J Vasc Surg* 1992; 15: 851–859.
3. Marty-Ane CH, Alric P, Picot MC, et al. Ruptured abdominal aortic aneurysm: influence of intraoperative management on surgical outcome. *J Vasc Surg* 1995; 22: 780–786.
4. Darling RC, Cordero JA, Chang BB. Advances in the surgical repair of ruptured abdominal aortic aneurysms. *Cardiovasc Surg* 1996; 4: 720–723.
5. Dardik A, Burleyson GP, Bowman H, et al. Surgical repair of ruptured abdominal aortic aneurysms in the state of Maryland: factors influencing outcome among 527 recent cases. *J Vasc Surg* 1998; 28: 413–421.
6. Noel AA, Gloviczki P, Cherry KJ Jr, et al. Ruptured abdominal aortic aneurysms: the excessive mortality rate of conventional repair. *J Vasc Surg* 2001; 34: 41–46.
7. Veith FJ. Emergency abdominal aortic aneurysm surgery. *Compr Ther* 1992; 18: 25–29.
8. Marin ML, Veith FJ, Cynamon J, et al. Initial experience with transluminally placed endovascular grafts for the treatment of complex vascular lesions. *Ann Surg* 1995; 222: 449–469.
9. Yusuf SW, Whitaker SC, Chuter TA, et al. Emergency endovascular repair of leaking aortic aneurysms. *Lancet* 1994; 344: 1645.
10. Parodi JC, Palmaz JC, Barone HD. Transfemoral intraluminal graft implantation for abdominal aortic aneurysms. *Ann Vasc Surg* 1991; 5: 491–499.
11. Andresen AFR. Results of treatment of massive gastric hemorrhage. *Am J Digest Dis* 1939; 6: 641–650.
12. Andresen AFR. Management of gastric hemorrhage. *NY State J Med* 1948; 48: 603–611.
13. Shaftan GW, Chiu CJ, Dennis C, Harris B. Fundamentals of physiologic control of arterial hemorrhage. *Surgery* 1965; 58: 851–856.
14. Bickell WH, Wall MJ Jr, Pepe PE, et al. Immediate versus delayed fluid resuscitation for hypotensive patients with penetrating torso injuries. *N Engl J Med* 1994; 331: 1105–1109.
15. Crawford ES. Ruptured abdominal aortic aneurysm: an editorial. *J Vasc Surg* 1991; 13: 348–350.
16. Mehta M, Darling RC III, Roddy SP, et al. Factors associated with abdominal compartment syndrome complicating endovascular repair of ruptured abdominal aortic aneurysms. *J Vasc Surg* 2005; 42: 1047–1051.
17. Veith FJ, Ohki T. Endovascular approaches to ruptured infrarenal aorto-iliac aneurysms. *J Cardiovasc Surg* 2002; 43: 369–378.
18. Yusuf SW, Whitaker SC, Chuter TA, et al. Early results of endovascular aortic aneurysm surgery with aortouniiliac graft, contralateral iliac occlusion, and femorofemoral bypass. *J Vasc Surg* 1997; 25: 165–172.
19. Yusuf SW, Hopkinson BR. Is it feasible to treat contained aortic aneurysm rupture by stent-graft combination? In: Greenhalgh RM, ed. *Indications in Vascular and Endovascular Surgery*. London, England: WB Saunders; 1998; 153–165.
20. Lachat ML, Pfammatter T, Witzke HJ, et al. Endovascular repair with bifurcated stent-grafts under local anaesthesia to improve outcome of ruptured aortoiliac aneurysms. *Eur J Vasc Endovasc Surg* 2002; 23: 528–536.
21. Ohki T, Veith FJ, Sanchez LA, et al. Endovascular graft repair of ruptured aorto-iliac aneurysms. *J Am Coll Surg* 1999; 189: 102–123.
22. Malina M, Veith F, Ivancev K, Sonesson B. Balloon occlusion of the aorta during endovascular repair of ruptured abdominal aortic aneurysm. *J Endovasc Ther* 2005; 12: 556–559.
23. Hughes CW. Use of an intra-aortic balloon catheter tamponade for controlling intra-abdominal hemorrhage in man. *Surgery* 1954; 36: 65–68.
24. Hesse FG, Kletschka HD. Rupture of abdominal aortic aneurysm: control of hemorrhage by intraluminal balloon tamponade. *Ann Surg* 1962; 155: 320–322.
25. Ng AC, Ochsner EC. Use of Fogarty catheter tamponade for ruptured abdominal aortic aneurysms. *Am J Roentgenol* 1977; 128: 31–33.
26. Hyde GL, Sullivan DM. Fogarty catheter tamponade of ruptured abdominal aortic aneurysms. *Surg Gynecol Obstet* 1982; 154: 197–199.

Thoracic Aortic Aneurysms and Dissections: Endovascular Treatment

Donald T. Baril, MD, Jae S. Cho, MD, Rabih A. Chaer, MD,
and Michel S. Makaroun, MD

Division of Vascular Surgery, University of Pittsburgh Medical Center, Pittsburgh, PA

OUTLINE

ENDOGRAFTS FOR THORACIC ENDOVASCULAR AORTIC REPAIR
- **Gore Tag Device**
- **Medtronic Talent Device**
- **Cook Zenith TX2 Device**

INDICATIONS FOR THORACIC ENDOVASCULAR AORTIC REPAIR

ANATOMIC CONSIDERATIONS FOR THORACIC ENDOVASCULAR AORTIC REPAIR
- **Iliac Anatomy**
- **Landing Zones**

PREOPERATIVE PLANNING
- **Imaging**
- **Sizing**
- **Landing Zone Evaluation**
- **Spinal Cord Protection**

RESULTS OF THORACIC ENDOVASCULAR AORTIC REPAIR
- **TAG Results**
- **The VALOR Trial**
- **The Cook Zenith TX1 and TX2 Thoracic Endograft Trial**

COMPLICATIONS OF THORACIC ENDOVASCULAR AORTIC REPAIR

ENDOVASCULAR TREATMENT OF TYPE B AORTIC DISSECTION

CONCLUSION

ABSTRACT

The treatment of thoracic aortic disease has changed radically with the advances made in endovascular therapy since the concept of thoracic endovascular aortic repair was first described 15 years ago. Currently, there is a diverse array of endografts that are commercially available to treat the thoracic aorta. Multiple studies, including industry-sponsored and single-institution reports, have demonstrated excellent outcomes of thoracic endovascular aortic repair for the treatment of thoracic aortic aneurysms, with less reported perioperative morbidity and mortality in comparison with conventional open repair. Additionally, similar outcomes have been demonstrated for the treatment of type B dissections. However, the technology remains relatively novel, and larger studies with longer term outcomes are necessary to more fully evaluate the role of endovascular therapy for the treatment of thoracic aortic disease. This review examines the currently available thoracic endografts, preoperative planning for thoracic endovascular aortic repair, and outcomes of thoracic endovascular aortic repair for the treatment of both thoracic aortic aneurysms and type B aortic dissections. ***Mt Sinai J Med 77:256–269, 2010.*** **©** *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Donald T. Baril
Division of Vascular Surgery
University of Pittsburgh
School of Medicine
Pittsburgh, PA
Email: barildt@upmc.edu

Key Words: aortic dissection, endovascular, thoracic aortic aneurysm, thoracic endovascular aortic repair.

Endovascular abdominal aortic aneurysm repair (EVAR) was first described by Parodi in 1991.[1] Soon following this, Dake introduced the same concept of stent graft exclusion to the thoracic aorta in 1994 for the treatment of thoracic aortic aneurysms (TAAs).[2] These initial stent grafts were custom-designed for each patient and were constructed of self-expanding stainless steel stents covered with woven Dacron grafts. Ten years later, thoracic endovascular aortic repair (TEVAR) subsequently became a mainstream approach to the treatment of TAAs with the release of commercial endografts.

Unlike EVAR, which has been compared with open surgical alternatives in randomized controlled trials, most data regarding the outcomes of TEVAR originate from industry-sponsored nonrandomized comparisons to open repair of TAA, in addition to single-center experiences and registry data that combine multiple different pathologies and indications. Despite this, TEVAR has emerged as an

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20178

© 2010 Mount Sinai School of Medicine

attractive option for the treatment of TAA with very good early and mid-term results. Furthermore, this same technology has been more recently applied to the treatment of type B aortic dissections.

> *Thoracic endovascular aortic repair (TEVAR) has emerged as an attractive option for the treatment of thoracic aortic aneurysms (TAAs) with very good early and mid-term results. Furthermore, this same technology has been more recently applied to the treatment of type B aortic dissections.*

ENDOGRAFTS FOR THORACIC ENDOVASCULAR AORTIC REPAIR

Since the initial description of TEVAR, a significant amount of progress has been made and continues toward the creation of the ideal endograft. In addition to the challenges faced with devices for EVAR, the endografts used for TEVAR need to address the significant hemodynamic forces in the thoracic aorta along with the proximity of the great vessels, while treating a larger diameter vessel. At present time, there are 3 commercially available devices that have been approved by the United States Food and Drug Administration for the treatment of TAAs. Additionally, there is another device (Relay, Bolton Medical Inc., Sunrise, FL), which is currently undergoing a phase II trial in the United States.

Gore Tag Device

The TAG endoprosthesis (W.L. Gore & Associates, Inc., Flagstaff, AZ) is made of an expanded polytetrafluoroethylene (ePTFE) tube reinforced with ePTFE/FEP (fluorinated ethylene propylene) film and an external nickel-titanium alloy (Nitinol) self-expanding stent affixed to the graft with ePTFE/FEP bonding tape. This device deploys from the middle of the graft toward each end, a design characteristic intended to prevent windsocking at the proximal fixation zone (Figure 1). A circumferential PTFE sealing cuff is located at the base of each flared, scalloped end. The device is constrained by an ePTFE tape that is released during deployment from the middle of the device toward both ends. Currently, devices are available in 26- to 40-mm diameters and 10-, 15-, or 20-cm lengths. The device profile requires 20- to 24-French (F) introducer sheaths. A 45-mm-diameter device is under clinical testing. A minimum of 2-cm neck is required for sealing on both ends, and a 7% to 18% device oversizing to the native aorta is recommended.

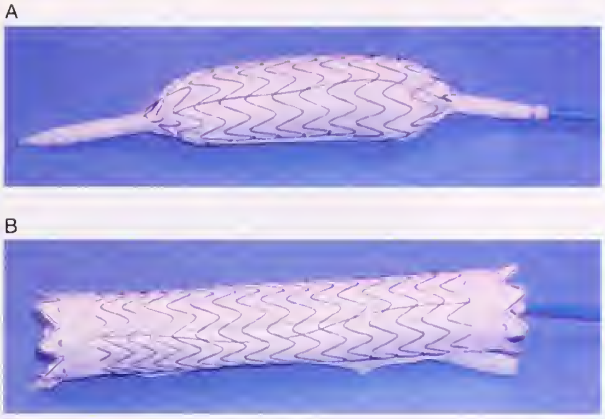


Fig 1. (A) Gore TAG endoprosthesis during deployment, opening from the middle toward each end. (B) Fully deployed Gore TAG endoprosthesis.

Medtronic Talent Device

The Talent device (Medtronic, Santa Rosa, CA) is composed of a thin woven polyester fabric sewn to a self-expanding Nitinol wire frame (Figure 2). The device comes in 4 different configurations: proximal main, proximal extension, distal main, and distal extension. The proximal stent grafts and the distal extension are equipped with a bare spring to improve fixation and allow deployment across the origins of the great vessels proximally and the celiac axis distally. It has the widest range of device diameters, ranging from 22 to 46 mm. Lengths range only from 112 to 116 mm. A 10% to 20% oversizing to the aortic neck is recommended, and a 2-cm sealing zone is required in the neck. Tapered grafts are also available to accommodate the diameter differences along the length of the aorta. The grafts are loaded in a delivery system with a 20- to 24-F size. The Valiant device is a modified Talent without a longitudinal bar, with increased stent peaks and closed web options (Figure 3). It is available in longer lengths of 22 cm and is currently undergoing clinical testing in the United States.[3]

Cook Zenith TX2 Device

The Zenith TX2 (Cook, Inc., Bloomington, IN) is a 2-piece modular graft made of woven Dacron

Fig 2. Medtronic Talent endograft.

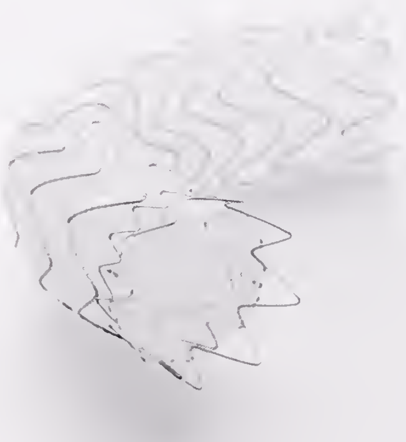


Fig 3. Medtronic Valiant endograft.

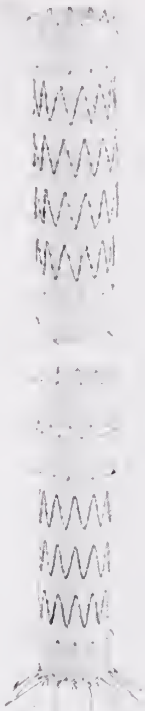


Fig 4. Cook TX2 endograft.

fabric sewn to self-expanding stainless steel z-stents (Figure 4). The stents are on the inside of the graft in the sealing zone but are on the outside in the rest of the endoprosthesis. The device is available in diameters from 28 to 42 mm and in lengths from 120 to 216 mm. Active fixation is provided by barbs on both ends, caudally oriented in the proximal sealing stent and cranially directed on a bare distal stent, specifically intended to prevent graft migration and component separation during conformational changes in the aneurysm after TEVAR. A minimum 3-cm landing zone is recommended. The device is contained in an introduction system of 20 or 22 F in diameter.

INDICATIONS FOR THORACIC ENDOVASCULAR AORTIC REPAIR

Indications to treat TAAs depend on the aneurysm size, rate of growth, location, symptoms, and general medical condition of the patient. Similar criteria should be applied for both open and endovascular repair and include a size of 6 cm or larger, saccular morphology, and symptomatic aneurysms including rupture (Figure 5). The size recommendations are variable, as no randomized trials exist to guide the decision-making process. Considerations in the choice of TEVAR over open repair are primarily anatomic in nature. An appropriate landing zone should be available both proximally and distally to allow for sealing and exclusion of the aneurysm from circulation as well as an appropriately sized arterial access to deliver the stent graft to its desired location. In addition to anatomy, selection of TEVAR should also include age and medical risk in the decision tree.

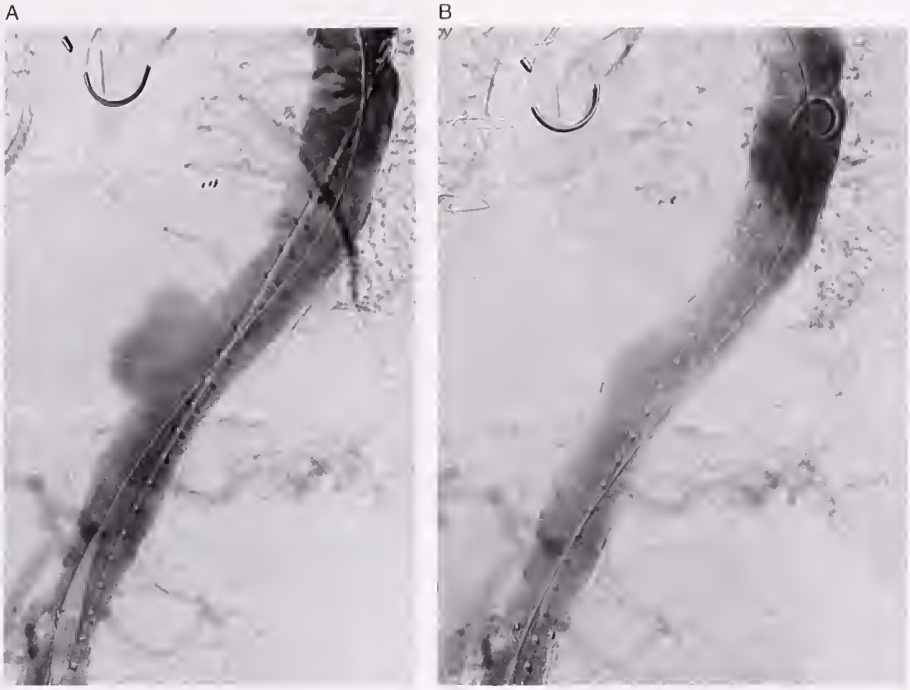


Fig 5. (A) Angiogram demonstrating saccular aneurysm of the descending thoracic aorta. (B) Angiogram following successful thoracic endovascular aortic repair (TEVAR) with preservation of the celiac axis.

Considerations in the choice of TEVAR over open repair are primarily anatomic in nature. An appropriate landing zone should be available both proximally and distally to allow for sealing and exclusion of the aneurysm from circulation as well as an appropriately sized arterial access to deliver the stent graft to its desired location. In addition to anatomy, selection of TEVAR should also include age and medical risk in the decision tree.

In the presence of suitable anatomy, TEVAR seems like the optimal choice in most patients due to its reported lower morbidity and mortality. However, the decision should be considered in the setting of the limited extended follow-up data of TEVAR and the requirement for lifelong follow-up. Consideration should be given to thoracotomy in young patients with a long life expectancy who are low risk for open repair. For high-risk individuals, clinical results favor TEVAR given the perioperative risks of open repair. As more long-term data become available, a more liberal use of TEVAR in lower risk individuals may become the standard, especially given the significant perioperative risk reduction.[4]

Unfortunately, most patients with TAAs lack ideal anatomy for TEVAR, and the risk-benefit assessment becomes more complex. TAAs often impinge on a major arterial branch that must be covered for adequate sealing. In these situations, the safety of such branch sacrifice must be balanced against further observation of the aneurysm or open repair. Alternatively, surgical bypass or debranching

Unfortunately, most patients with TAAs lack ideal anatomy for TEVAR and the risk-benefit assessment becomes more complex. TAAs often impinge on a major arterial branch that must be covered for adequate sealing. In these situations the safety of such branch sacrifice must be balanced against further observation of the aneurysm or open repair.

to maintain perfusion while extending the sealing zone may be an option for select patients.

There are a number of relative contraindications to TEVAR. In particular, TEVAR should be contraindicated when the anatomy of the aneurysm prevents the safe and effective performance of the endograft, often increasing the short-term and long-term complication rates. Aortic size outside the treatment range of available devices should be avoided. Severe tortuosity of the aneurysm or the access vessels can also increase morbidity. Advanced aortoiliac occlusive disease and circumferential thrombus at the attachment sites all preclude the safe performance of TEVAR and should be avoided. TEVAR should also be avoided in patients with severe allergies to nickel, in patients with significant chronic renal insufficiency, and in patients who cannot be followed reliably.

ANATOMIC CONSIDERATIONS FOR THORACIC ENDOVASCULAR AORTIC REPAIR

Preoperative planning and assessment of the anatomy are paramount to successful TEVAR. Specific anatomic requirements differ among the various endografts, but adequate arterial access and sealing zones are essential for all patients undergoing TEVAR.

Iliac Anatomy

The common femoral arteries are most commonly utilized to access the aorta. The right iliac artery normally provides less of an angle at the aortic junction and, hence, the right is the favored access side. The left common femoral artery is most commonly used for angiography catheters to evaluate before, during, and after endograft deployment. Ideal access vessels have a lumen of at least 8 mm to allow for the introduction of the larger device sizes. The smallest diameter of the access vessels is typically in the proximal external iliac artery, an area that is also more prone to calcified plaques. If the narrowest segment of the access vessel is below the minimum required size for the planned device, access may be obtained more proximally in the common iliac artery by using an iliac conduit. In unusual circumstances where both iliac systems are severely diseased, consideration can be given to an access through the distal aorta.

Landing Zones

Optimal fixation of an endograft to both proximal and distal landing zones is essential to successful TEVAR. Ishimaru classified the thoracic aorta into 5 landing zones.[5] Each zone is bordered by a tangential line aligned with the distal sides of each great vessel; zone 0 involves the origin of the innominate artery, zone 1 the origin of the left common carotid artery (LCCA), zone 2 the origin of the left subclavian artery (LSA), zone 3 the proximal descending thoracic aorta down to the T4 vertebral body, and zone 4 the remainder of descending thoracic aorta. TAAs normally do not involve one focal segment of the aorta, and typically the degenerative process involves the adjacent aorta. Although 2 to 3 cm is considered an adequate landing zone for most devices, a longer seal zone may be preferable, particularly at angulated areas, to decrease the occurrence of late endoleaks. However, in increasing the length of coverage, consideration needs to be made to the increasing risk of spinal cord ischemia (SCI) with the sacrifice of multiple intercostal and lumbar arteries. The proximity of the sealing zones to major branches should be carefully evaluated to avoid occlusion of a necessary artery or, alternatively, to plan a bypass to maintain patency of the branch.

PREOPERATIVE PLANNING

Imaging

Preoperative imaging is used to delineate the anatomy of the aneurysm, its relationship to adjacent arch and mesenteric vessels, and the access vessels. Computed tomographic angiography (CTA) is the primary modality for providing all of this relevant preoperative anatomic information.

Preoperative imaging is used to delineate the anatomy of the aneurysm, its relationship to adjacent arch and mesenteric vessels, as well as the access vessels. Computed tomographic angiography (CTA) is the primary modality for providing all of this relevant preoperative anatomic information.

CTA using 64-slice imaging allows for three-dimensional (3D) volume rendering, as well as maximum-intensity projections (MIPs) and sagittal and coronal reconstructions, which aid in the analysis of angles and relationships to adjacent vessels (Figure 6). Software companies such as M2S

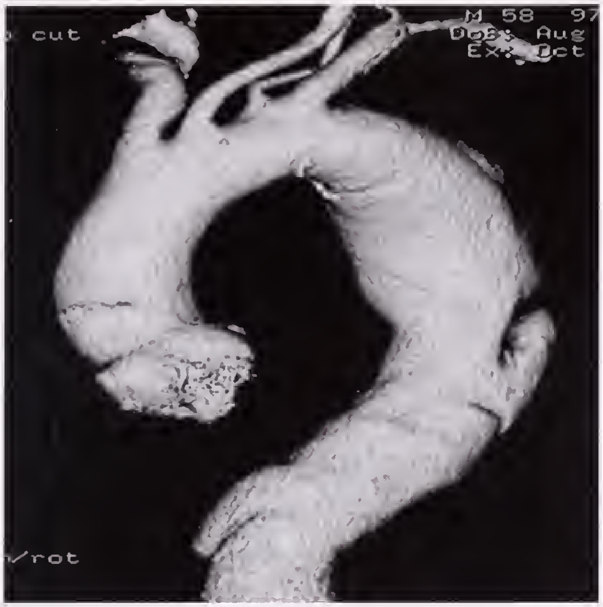


Fig 6. Three-dimensional reconstruction of a CT angiogram demonstrating thoracic aortic aneurysm (TAA) with focal dissection.

(Medical Metrx Software, West Lebanon, NH), Vitrea (Vital Images, Minnetonka, MN), and Aquarius (TeraRecon Inc, San Mateo, CA) offer additional tools such as reconstructions orthogonal to the line of flow and virtual grafts to aid in measurements and procedure planning. Magnetic resonance angiography (MRA) can be substituted for CTA in patients with significant renal failure.

The preoperative CTA should include the chest, abdomen, and pelvis, with and without contrast enhancement, to fully evaluate the access vessels as well as the extent of calcification and angulation throughout. Evaluation of arch disease for atheromatous plaque as well as the vertebral arteries is necessary to minimize the risk of cerebral ischemic complications. Additionally, determination of internal iliac artery flow can be critical in assessing perfusion to the spinal cord, especially in patients in whom a prior aortic reconstruction exists or when anticipating coverage of a long segment of aorta.

Sizing

Accurate diameter measurements of the access vessels and the sealing zones are important to determine the feasibility of the procedure and appropriate graft sizes. An undersized endograft can result in an endoleak or migration, while excessive oversizing may cause infolding, poor sealing and excessive radial force with accelerated degeneration of the neck.

Accurate diameter measurements of the access vessels and the sealing zones are important to determine the feasibility of the procedure and appropriate graft sizes. An undersized endograft can result in an endoleak or migration, while excessive oversizing may cause infolding, poor sealing, and excessive radial force with accelerated degeneration of the neck.

Aortic diameters are best measured on reconstructions orthogonal to the line of flow, especially in areas of tortuosity and angulation, to avoid incorrect measurements. Neck diameters can vary extensively between the proximal and distal sealing zones, which may lead to inappropriate oversizing on the smaller end. To minimize the impact of this phenomenon, tapered grafts may be used or, alternatively, shorter devices with varying diameters.

Appropriate length measurements are also necessary to estimate the required device lengths. The thoracic aorta is quite tortuous at baseline compared to the abdominal aorta, a quality that becomes exacerbated in the presence of an aneurysm (Figure 7). To best estimate length, a centerline flow measurement from a 3D reconstruction should be used, although this may underestimate the length of endografts needed, as the graft will tend to gravitate toward the greater curvature.

Landing Zone Evaluation

The lack of adequate landing zones both proximally and distally is encountered frequently. Extension into zone 2 to achieve adequate seal proximally was required at least 20% of the time in some regulatory trials.[6] Management of the left subclavian artery with either coverage or bypass has undergone a substantial change since the early period of TEVAR. Initially, all cases that necessitated left subclavian artery coverage underwent a left carotid-subclavian bypass or transposition until reports of safe coverage surfaced resulting in indiscriminate coverage of the left subclavian artery.[7–11] However, a resurgence of interest in preserving flow in the left subclavian artery has developed following reports of cases of brainstem stroke, an increased risk of SCI, or severe arm ischemia.[12,13]

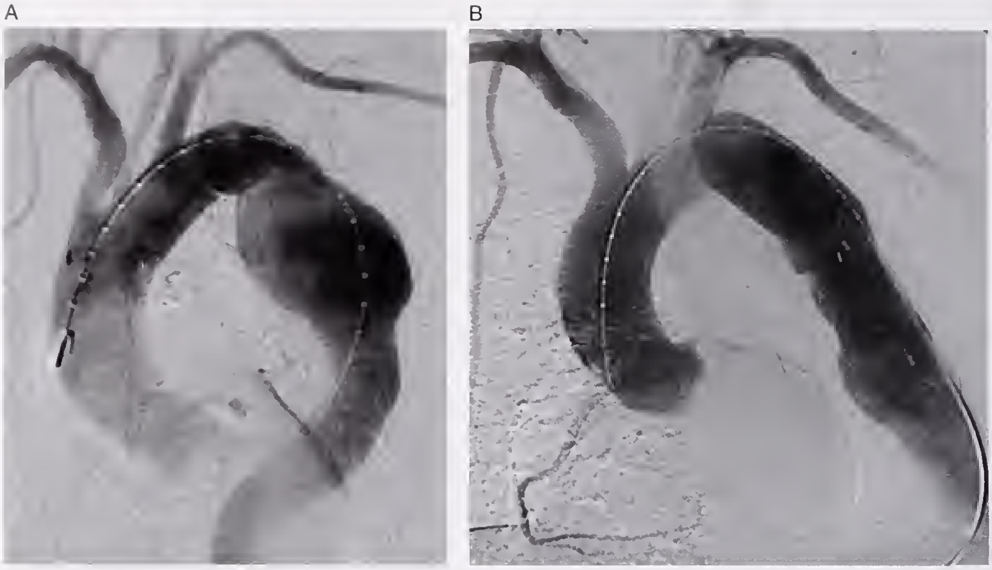


Fig 7. (A) Tortuous descending thoracic aortic aneurysm (TAA). (B) Descending thoracic aorta with complete aneurysm exclusion following thoracic endovascular aortic repair (TEVAR).

Prior to coverage of the left subclavian artery, preoperative imaging should be reviewed to assess the patency of the right vertebral artery into the basilar as well as the circle of Willis so as to identify patients who cannot tolerate left subclavian artery occlusion. Mandatory left subclavian artery revascularization is indicated in the presence of a dominant left vertebral artery, a patent left internal mammary bypass to a coronary artery, an aberrant right subclavian artery, a hypoplastic right vertebral artery, or an anomalous origin of the left vertebral artery from the arch.[12,14,15] Additionally, revascularization of the left subclavian artery should be considered when an increased risk of paraplegia exists with extensive coverage of the thoracic aorta combined with a previous infrarenal aortic replacement.[12,16] Delayed revascularization can usually be performed in cases of arm ischemia without untoward effects.[12,15]

When the distal landing zone is inadequate, the operator faces similar options as with the left subclavian artery. When adequate collateral circulation between the celiac and superior mesenteric artery (SMA) primarily through the gastroduodenal artery is established by preoperative angiography, the celiac axis may be covered safely to extend the distal landing zone during TEVAR.[16–18] If the collateral network is not adequate, a bypass to the celiac circulation must be performed prior to coverage.

Spinal Cord Protection

Unfortunately, TEVAR is not without the risk of spinal cord ischemic complications; patients with long segment aortic coverage and previous abdominal aortic grafting are at particularly high risk.[19–21] For such patients, particular preoperative planning and use of adjunctive therapy should be utilized to help minimize this risk.

> *Unfortunately, TEVAR is not without the risk of spinal cord ischemic complications; patients with long segment aortic coverage and previous abdominal aortic grafting are at particularly high risk.*

The most commonly used of these adjuncts is cerebrospinal fluid (CSF) drainage. The drain is typically inserted preoperatively and maintained for approximately 48 hours, keeping the pressure <10 cm H_2O. The incidence of hemorrhagic complications from catheter placement is also not negligible and reported from 0% to 3%.[22] Whether used preoperatively or not, CSF drainage has been shown to be effective in reversal of delayed-onset paraplegia after TEVAR and should be employed whenever SCI is suspected.[21]

Hemodynamic stability also appears to play a significant role in spinal cord protection, as the underlying etiology appears to be poor collateral flow and decreased perfusion of the spinal cord.

A low mean intraoperative arterial blood pressure of ≤ 70 mm Hg has been linked to SCI in one review.[21] Except for the short time required for endograft deployment when a relatively low blood pressure may be beneficial, a rather elevated pressure is preferred intraoperatively. Additionally, postoperative hypotension has been noted to precipitate SCI events in patients after TEVAR.

Naloxone has been the most-studied pharmacologic adjunctive therapy. Intravenous naloxone has been shown to be effective in decreasing circulating levels of excitotoxic amino acids and antagonizing opiate-mediated spinal cord vasoconstriction following open TAA repair. Its beneficial effects have been observed in open procedures[23,24] and it may be utilized in selected cases of TEVAR.

RESULTS OF THORACIC ENDOVASCULAR AORTIC REPAIR

Although there has been a significant amount of data presented on the outcomes of TEVAR, there continues to be a paucity of randomized and long-term outcome information. With regard to the single-center and registry data, the patient populations and anatomic locations are quite diverse, making interpretation of results somewhat difficult. The majority of the most applicable data has been derived from the trials of the 3 FDA-approved devices.[25–27] In these trials, the patients who underwent TEVAR were deemed good-risk surgical candidates who had elective procedures.

Although there has been a significant amount of data presented on the outcomes of TEVAR, there continues to be a paucity of randomized and long-term outcome information. The majority of the most applicable data has been derived from the trials of the 3 FDA-approved devices. In these trials, the patients who underwent TEVAR were deemed good-risk surgical candidates who had elective procedures.

TAG Results

The Gore TAG thoracic endoprosthesis was evaluated in a phase II multicenter trial.[25] This was a prospective, nonrandomized, controlled multicenter trial. The trial enrolled 140 study patients and 94 control subjects between September 1999 and May 2001 through 17 clinical sites in the United States. The control group consisted of 44 patients acquired prospectively during the study and 50 historical patients acquired by selecting the most recent surgical patients in reverse chronological order.

The Gore TAG thoracic endoprosthesis was evaluated in a prospective, nonrandomized, controlled multicenter phase II trial. Of 142 patients recruited (140 in the pivotal trial and 2 extended access), 139 (98%) underwent successful implantation of the TAG device. The 3 failures were all due to poor iliac access.

A conduit was required to facilitate access in 21 patients (15%). Prophylactic left carotid/subclavian bypass grafting was performed in 28 patients in preparation for planned left subclavian artery coverage with the device. The mean descending thoracic aortic aneurysm (DTA) size was 64.1 ± 15.4 mm. Ninety percent were American Society of Anesthesiologists category III or IV. One device was used in 44% of patients, while 56% of patients required 2 or more devices to bridge the thoracic aorta. The procedure time averaged 150 minutes, estimated blood loss averaged 506 mL, ICU stay averaged 2.6 days, and hospital stay averaged 7.6 days. Within 30 days, 45 (32%) patients had at least 1 major adverse event: 5 (4%) experienced a stroke, 4 (3%) demonstrated temporary or permanent paraplegia, 20 (14%) experienced vascular trauma or thrombosis, and 2 (1.5%) died. Mean follow-up was 24.0 months. Four patients had aneurysm-related deaths. Three patients underwent endovascular revisions for endoleak. No ruptures were reported. Twenty wire fractures were identified in 19 patients; 18 (90%) of these occurred in the longitudinal spine and only 1 patient required treatment. At 2 years, aneurysm-related and overall survival rates were 97% and 75%, respectively.

The VALOR Trial

The VALOR trial (Vascular Talent Thoracic Stent Graft System for the Treatment of Thoracic Aortic Aneurysms) was a prospective, multicenter study with 3 arms conducted at 38 sites in the U.S.[26] The trial enrolled 195 patients between December 2003 and June 2005. Results were compared with retrospective open surgical data of 189 patients from 3 centers of excellence in thoracic aortic surgery.

Compared with the open surgery group, the VALOR test group had similar age and sex distributions, but had a smaller TAA size. Patients undergoing TEVAR received a mean number of 2.7 endograft components. The diameters of 25% of the proximal stent graft components implanted were <26 mm or >40 mm. Revascularization of the left subclavian artery was performed prior to the initial TEVAR procedure in 5.2% of patients. Twenty-one percent of patients required iliac conduits. The bare spring segment of the most proximally implanted device was in zones 1 or 2 of the aortic arch in 33.5% of patients. Vessel access and endograft deployment were successful at the intended site for 194 patients (99.5%). The 30-day perioperative mortality was 2.1%.

> *The VALOR trial (Vascular Talent Thoracic Stent Graft System for the Treatment of Thoracic Aortic Aneurysms) was a prospective, multicenter study with 3 arms conducted at 38 sites in the United States. The trial enrolled 195 patients, whose results were compared with retrospective open surgical data of 189 patients from three centers of excellence in thoracic aortic surgery. Vessel access and endograft deployment were successful at the intended site for 194 trial patients (99.5%). The 30-day perioperative mortality was 2.1%.*

Perioperative paraparesis was 7.2%, paraplegia was 1.5%, and stroke was 3.6%. Overall major adverse-event rate was 41%. Twelve-month all-cause mortality was 16.1%, with an aneurysm-related mortality rate of 3.1%. Twelve-month endoleak rate was 12.2%. Conversion to open surgery was necessary in 0.5%. Stable or decreasing aneurysm diameter was seen in 91.5%. Compared with the historical open control group, TEVAR had fewer 30-day major adverse events (41% versus 84.4%, $P < .001$), lower perioperative mortality (2% versus 8%, $P < .01$), and lower 12-month aneurysm-related mortality (3.1% versus 11.6%, $P < .002$).

The Cook Zenith TX1 and TX2 Thoracic Endograft Trial

This was a North American multicenter, nonrandomized, prospective clinical trial designed to assess the technical success and outcome after patients with thoracic aortic pathology, at high risk for conventional therapy, who were treated with the Zenith TX1 and TX2 endovascular grafts.[27] Between 2001 and 2004, patients at high risk for conventional surgical therapy presenting with chronic aortic dissections, thoracic aneurysms, or aortobronchial or aortoesophageal fistulas were treated with a single- or multiple-piece endovascular graft. In total 100 patients were treated, including 81 aneurysms, 15 aortic dissections (with aneurysms), 2 patients with fistulous connections (1 aortobronchial and 1 aortoesophageal), 1 subclavian artery aneurysm, and 1 aortic rupture. Mean follow-up was 14 months and mean aneurysm size 62 mm. The majority of patients (55%) had undergone prior aortic aneurysm repair. Surgical modifications were required to create adequate implantation sites in 29% of the patients, including 14 elephant trunk/arch reconstructions, 18 carotid-subclavian bypasses, and 4 visceral artery bypasses.

> *In the Cook Zenith TX1 and TX2 Thoracic Endograft Trial, 100 patients presenting with chronic aortic dissections, thoracic aneurysms, or aortobronchial or aortoesophageal fistulas and at high risk for conventional surgical therapy were treated with a single- or multiple-piece endovascular graft. The majority of patients (55%) had undergone prior aortic aneurysm repair. Overall mortality was 17% and aneurysm-related mortality was 14% at 1 year.*

Iliac conduits were required in 19 patients. Overall mortality was 17% and aneurysm-related mortality was 14% at 1 year. SCI was suspected in 6 patients, and complete and permanent paraplegia developed in 2. Sac regression (>5 mm decrease in maximum diameter) was observed in 52% and 56% at 12 and 24 months, respectively. Growth was noted in 1 patient (1.6%) at 12 months. Endoleaks were detected in 8 patients (8.5%) at 30 days and in 3 patients (6%) at 12 months. Secondary interventions were required in 15 patients. Migration (>10 mm) of the proximal or distal stent was noted in 3 patients (6%). Two had proximal and 1 had distal migration, none of whom required treatment.

COMPLICATIONS OF THORACIC ENDOVASCULAR AORTIC REPAIR

As aforementioned, the overall rate of perioperative complications following TEVAR is less than that of open repair.[25–27] Not unexpectedly, cardiopulmonary complications dominate the early morbidity, although frequency and severity are less than for open repair.[26,28,29] Over the long term, the benefit of TEVAR with respect to occurrence of major adverse events appears to persist.[30]

> *The overall rate of perioperative complications following TEVAR is less than that of open repair. Not unexpectedly, cardiopulmonary complications dominate the early morbidity, although frequency and severity are less than for open repair.*

There are TEVAR-specific complications, including vascular complications, related for the most part to the use of large sheaths in diseased, atherosclerotic arteries. These were relatively high in the early TAG trial (14%).[25] The increased awareness of access issues has resulted in a reduction of vascular complications to the 5% to 10% range in most recent reports.[6,26,31] Serious iliac artery injuries, a subset of vascular complications ranging from a tear to complete avulsion and eversion of the sheath on withdrawal, are uncommon but potentially life threatening. The prophylactic use of iliac conduits should be encouraged to avoid these injuries.

Neurologic complications can be devastating after TEVAR. Although stroke receives less attention than SCI, it is not uncommon and occurs at similar rates between open and endovascular repairs, ranging from 3% to 5%.[6,12,14,26]

> *Neurologic complications of TEVAR can be devastating. Although stroke receives less attention than SCI, it is not uncommon and occurs at similar rates between open and endovascular repairs, ranging from 3% to 5%.*

Deployment of the proximal end of the endograft in zone 2 proximal to the left subclavian artery shows a strong association with perioperative stroke, likely secondary to manipulation of the arch with catheters, wires, balloons, and stent grafts at the origin of cerebral vessels.[6,18] The importance of operator's technique with minimal manipulation, meticulous cleansing of guidewires and catheters, and limited use of ballooning only inside the grafts cannot be overemphasized in minimizing the risk of stroke during TEVAR. Other factors that have been identified in the EUROSTAR (European Collaborators on Stent/Graft Techniques for Aortic Aneurysm Repair) trial as independent predictors of stroke included female sex and prolonged procedure time.[13]

As was initially observed with the TAG study that showed a marked reduction of paraplegia/paraparesis compared with open procedures (3% versus 14%), the advantage of spinal cord protection with TEVAR was again evident when the data from all 3 FDA trials were compared with open controls.[25–27]

> *The TAG study showed that, compared with open procedures, TEVAR was associated with a marked reduction of paraplegia/paraparesis (14% versus 3%). The same advantage in spinal cord protection was also evident in data pooled from all 3 FDA trials.*

The reduced incidence of SCI with TEVAR may be related to avoidance of thoracic aortic clamping and reduced hemodynamic disturbances at the spinal cord level due to lack of blood loss from intercostal backbleeding. It should also be noted that SCI is not confined to the perioperative period. Delayed recurrent SCI and paraplegia after TEVAR have been reported.[20,32,33]

As with EVAR, endoleaks may also complicate TEVAR. At 1 year, endoleaks were present in nearly 4% of TAG and TX2 patients and 12% of TALENT patients.[25–27] During long-term follow-up of the TAG study, approximately 4% of patients continued to have an endoleak at each yearly follow-up, up to 5 years.[30]

> *TEVAR may also be associated with endoleaks. At 1 year, endoleaks were present in nearly 4% of TAG and TX2 patients and 12% of TALENT patients. During long-term follow-up of the TAG study, approximately 4% of patients continued to have an endoleak at each yearly follow-up, up to 5 years.*

The incidence of endoleaks is decreasing with second-generation devices with design improvements and better conformability.[31] Endoleaks are in general less common than comparable EVAR series, with more type I and III endoleaks and less type II, and accounted for most but not all reinterventions.[34]

Sac enlargement rate of ≥5 mm is encountered more frequently with TEVAR than with abdominal EVAR and has been reported in 7% to 14% of patients at 1 year, with little device variations.[25–27] Only the TAG pivotal trial has so far reported late results in a longitudinal cohort, and sac enlargement was noted in nearly 20% at 5 years.[30] Shrinkage was observed in 50% and a stable sac was noted in the remaining patients. These rates may or may not apply to other grafts, as the ePTFE material in the TAG study graft was somewhat porous with a sac behavior that resembled the original abdominal Gore Excluder.[35] Both TAG and Excluder endografts were modified to a less-porous ePTFE material, resulting in a much higher rate of sac shrinkage in the infrarenal position.[36] Similarly, the data from the confirmatory trial with the modified TAG device using the low-permeability ePTFE has so far shown a lower incidence of sac enlargement at 2 years of 2.9%, compared with 12.9% with the original device ($P = 0.11$).

Late migration of endografts has been reported to occur in 0% to 30% of patients.[37,38] Current commercial devices tested in the United States pivotal trials seem to have a low migration potential, ranging from 0.7% to 3.9%. At 5 years, the TAG trial shows no late migration of >10 mm by investigator reporting.[30] Migration was noted on both ends of the devices, with the TX2 trial reporting 1 cranial and 2 caudal migrations through 1 year. None of these was associated with endoleak or sac expansion or required an intervention.

Current devices suffer very few device integrity issues. The original TAG device showed fractures in the longitudinal wire in 14% of patients during 5 years of follow-up, with only one patient suspected of a type III endoleak receiving an endovascular repair.[25] No stent fracture, barb separation, stent-to-graft separation or component separation was observed in the TX2 device in 1 year of follow-up.[27] Three Talent devices were noted to have a fracture during follow-up, 1 in the connecting bar.[26] One device was noted to have a fabric abrasion after explantation for rupture.

Early reinterventions have been noted in only 2.1% in the pivotal TAG cohort during the first year but in 10.7% of the Talent study.[25,26] The majority of reinterventions are for the treatment of endoleaks and are typically endovascular procedures. TX2 reintervention rates through 1 year were similar between the 2 groups (4.4% TEVAR versus 5.7% open).[27] At long term, the data available from the TAG study showed a total reintervention rate of 3.6% of patients treated.[30] Two patients had a surgical conversion, 1 for an aortoesophageal fistula and the other for an arch aneurysm and migration of the proximal device in a very large neck that did not meet inclusion and exclusion criteria.

ENDOVASCULAR TREATMENT OF TYPE B AORTIC DISSECTION

Uncomplicated type B aortic dissections are treated medically in the majority of patients, and excellent results can be expected when therapy is continued long-term. Acute operations are performed only when complications are present, such as rupture or malperfusion.[39,40] Other relative indications for intervention include refractory hypertension, persistent pain, and extension of the dissection flap. However, in select centers, there has been a trend toward early operation, including the placement of endografts over the primary tear in younger patients in an attempt to prevent the adverse late sequelae of persistent dissection.[41,42] Although this approach seems logical, excellent short-term and long-term results can be expected with aggressive medical therapy alone, with most patients never requiring early or late surgical intervention.[43]

Soon after Dake first described TEVAR for aneurysmal disease, he and his colleagues subsequently reported their results with TEVAR for the treatment of patients with type B dissections.[44] In this initial report, Dake treated four patients with Stanford type A dissections and 15 patients with Stanford type B dissections with a 16% early mortality rate and no subsequent deaths or aortic rupture during the 13-month follow-up period. Additionally, they reported a 79% rate of false lumen thrombosis. Subsequent to this initial report, Nienaber compared the results of TEVAR in 12 consecutive patients to open surgical repair in 12 matched controls.[45] Open surgery was associated with a 1-year mortality rate of 33% and serious adverse events in 42% of patients within 12 months. Over the same time period, there was no significant morbidity or mortality in the TEVAR group.

Following these initial reports, numerous other studies have been published reporting on the results of TEVAR for the treatment of type B dissection. At present, there are numerous ongoing studies, including trials comparing TEVAR with conventional open repair. Recently, a meta-analysis of 37 studies was published examining the results of endovascular stent grafting for acute type B dissections in 184 patients.[41] Using the International Registry of Acute Aortic Dissections (IRAD) database, 30-day survival with endovascular therapy compared favorably with medical therapy, both of which were significantly better than open surgical therapy. In 2006, the IRAD investigators reviewed 242 patients with acute dissection undergoing either medical (78%), open surgical (11%), or endovascular (11%) therapy with an overall hospital mortality rate of 12%.[46] Survival at 3 years was similar in all groups: 78% ± 7% for medical therapy, 83% ± 19% for open surgical therapy, and 76% ± 25% for endovascular therapy.

A meta-analysis of 37 studies involving 184 patients treated for acute type B dissections with endovascular stent grafting showed that 30-day survival with endovascular therapy compared favorably with medical therapy, both of which were significantly better than open surgical therapy. A subsequent review involving 242 patients treated for acute dissection with medical, open surgical, and endovascular therapy showed similar 3-year survival rates for all three as follows: medical therapy, 78% ± 7%; open surgery, 83% ± 19%; and endovascular therapy, 76% ± 25%.

Less is known about the efficacy of TEVAR for the treatment of chronic type B dissections. Many believe that this remains the least favorable indication for TEVAR secondary to fibrotic changes of the intimal flap and aortic wall, which can limit expansion of the true lumen following endograft placement. Regardless, the most common indication for surgical intervention on the aorta itself in chronic type B dissections is early growth (>5-cm diameter) or recurrent refractory pain in the early months following successful initial medical therapy. Subacute dilation is usually isolated to the proximal descending aorta, for which elective placement of an endograft over the primary tear may be considered if there are adequate proximal and distal landing zones.[47] The goal is not to remove all dissected aorta but rather to address the area of maximal dilation (Figure 8). However, this may risk creating intimal tears at the distal landing zone with potential for dissection extension. Unfortunately, the ability to predict intimal flap movement and subsequent true lumen expansion is limited. As with the use of TEVAR for the treatment of TAAs, it is apparent that larger studies with longer follow-up are necessary to more completely evaluate TEVAR as a treatment modality for both acute and chronic type B dissection.

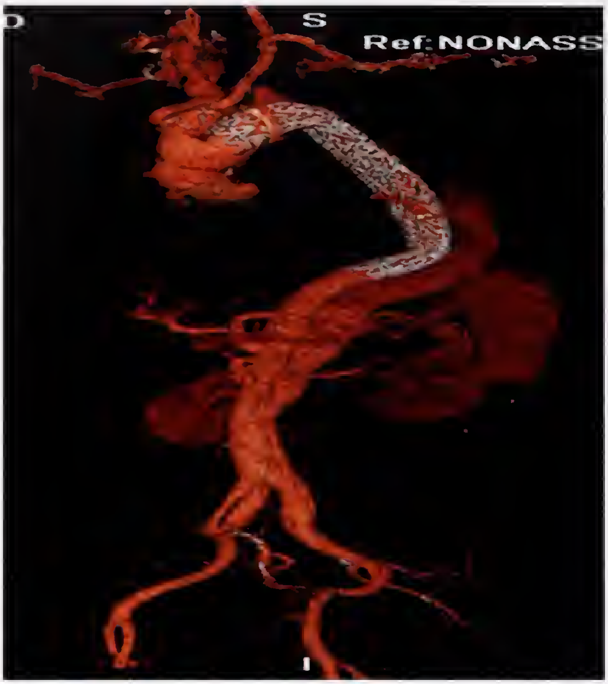


Fig 8. Three-dimensional reconstruction of a CT angiogram following thoracic endovascular aortic repair (TEVAR) for the treatment of aneurysmal degeneration of a chronic type B dissection.

CONCLUSION

The treatment of thoracic aortic disease has changed radically with the advances made in endovascular therapy since the concept of TEVAR was first described 15 years ago. There is a diverse array of endografts that are commercially available to treat the thoracic aorta. Furthermore, multiple studies have demonstrated excellent outcomes of TEVAR for the treatment of TAAs, with less reported perioperative morbidity and mortality in comparison with with conventional open repair. More recently, similar outcomes have been demonstrated for the treatment of type B dissections. However, the technology remains relatively novel, and larger studies with longer term outcomes are necessary to more fully evaluate the role of endovascular therapy for the treatment of thoracic aortic disease. Regardless, it is clear that the devices and the associated techniques will continue to advance to improve outcomes in the coming years.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Parodi JC, Palmaz JC, Barone HD. Transfemoral intraluminal graft implantation for abdominal aortic aneurysms. *Ann Vasc Surg* 1991; 5: 491–499.
2. Dake MD, Miller DC, Semba CP, et al. Transluminal placement of endovascular stent-grafts for the treatment of descending thoracic aortic aneurysms. *N Engl J Med* 1994; 331: 1729–1734.
3. Brooks M, Loftus I, Morgan R, Thompson M. The Valiant thoracic endograft. *J Cardiovasc Surg (Torino)* 2006; 47: 269–278.
4. Makaroun MS, Dillavou ED, Wheatley GH, Cambria RA. Five-year results of endovascular treatment with the Gore TAG device compared to open repair of thoracic aortic aneurysms. *J Vasc Surg* 2008; 47: 912–918.
5. Mitchell RS, Ishimaru S, Ehrlich MP, et al. First International Summit on Thoracic Aortic Endografting: roundtable on thoracic aortic dissection as an indication for endografting. *J Endovasc Ther* 2002; 9(suppl 2): II98–II105.
6. Cho JS, Haider SE, Makaroun MS. US multicenter trials of endoprostheses for the endovascular treatment of descending thoracic aneurysms. *J Vasc Surg* 2006; 43(suppl A): 12A–19A.
7. Sunder-Plassmann L, Scharrer-Pamler R, Liewald F, et al. Endovascular exclusion of thoracic aortic aneurysms: mid-term results of elective treatment and in contained rupture. *J Card Surg* 2003; 18: 367–374.
8. Rehders TC, Petzsch M, Ince H, et al. Intentional occlusion of the left subclavian artery during stent-graft implantation in the thoracic aorta: risk and relevance. *J Endovasc Ther* 2004; 11: 659–666.
9. Caronno R, Piffaretti G, Tozzi M, et al. Intentional coverage of the left subclavian artery during endovascular stent graft repair for thoracic aortic disease. *Surg Endosc* 2006; 20: 915–918.
10. Thompson M, Ivaz S, Cheshire N, et al. Early results of endovascular treatment of the thoracic aorta using the Valiant endograft. *Cardiovasc Intervent Radiol* 2007; 30: 1130–1138.
11. Burks JA Jr., Faries PL, Gravereaux EC, et al. Endovascular repair of thoracic aortic aneurysms: stent-graft fixation across the aortic arch vessels. *Ann Vasc Surg* 2002; 16: 24–28.
12. Peterson BG, Eskandari MK, Gleason TG, Morasch MD. Utility of left subclavian artery revascularization in association with endoluminal repair of acute and chronic thoracic aortic pathology. *J Vasc Surg* 2006; 43: 433–439.
13. Buth J, Harris PL, Hobo R, et al. Neurologic complications associated with endovascular repair of thoracic aortic pathology: incidence and risk factors. A study from the European Collaborators on Stent/Graft Techniques for Aortic Aneurysm Repair (EUROSTAR) registry. *J Vasc Surg* 2007; 46: 1103–1110.
14. Riesenman PJ, Farber MA, Mendes RR, et al. Coverage of the left subclavian artery during thoracic endovascular aortic repair. *J Vasc Surg* 2007; 45: 90–94.
15. Reece TB, Gazoni LM, Cherry KJ, et al. Reevaluating the need for left subclavian artery revascularization with thoracic endovascular aortic repair. *Ann Thorac Surg* 2007; 84: 1201–1205.
16. Sunder-Plassmann L, Orend KH. Stentgrafting of the thoracic aorta-complications. *J Cardiovasc Surg (Torino)* 2005; 46: 121–130.

17. Vaddineni SK, Taylor SM, Patterson MA, Jordan WD Jr. Outcome after celiac artery coverage during endovascular thoracic aortic aneurysm repair: preliminary results. *J Vasc Surg* 2007; 45: 467–471.
18. Criado FJ, bul-Khoudoud OR, Domer GS, et al. Endovascular repair of the thoracic aorta: lessons learned. *Ann Thorac Surg* 2005; 80: 857–863.
19. Gravereaux EC, Faries PL, Burks JA, et al. Risk of spinal cord ischemia after endograft repair of thoracic aortic aneurysms. *J Vasc Surg* 2001; 34: 997–1003.
20. Baril DT, Carroccio A, Ellozy SH, et al. Endovascular thoracic aortic repair and previous or concomitant abdominal aortic repair: is the increased risk of spinal cord ischemia real? *Ann Vasc Surg* 2006; 20: 188–194.
21. Chiesa R, Melissano G, Marrocco-Trischitta MM, et al. Spinal cord ischemia after elective stent-graft repair of the thoracic aorta. *J Vasc Surg* 2005; 42: 11–17.
22. Weaver KD, Wiseman DB, Farber M, et al. Complications of lumbar drainage after thoracoabdominal aortic aneurysm repair. *J Vasc Surg* 2001; 34: 623–627.
23. Acher CW. Multifactoral nature of spinal cord circulation. *Semin Thorac Cardiovasc Surg* 1998; 10: 7–10.
24. Acher CW, Wynn MM, Hoch JR, et al. Combined use of cerebral spinal fluid drainage and naloxone reduces the risk of paraplegia in thoracoabdominal aneurysm repair. *J Vasc Surg* 1994; 19: 236–246.
25. Makaroun MS, Dillavou ED, Kee ST, et al. Endovascular treatment of thoracic aortic aneurysms: results of the phase II multicenter trial of the GORE TAG thoracic endoprosthesis. *J Vasc Surg* 2005; 41: 1–9.
26. Fairman RM, Criado F, Farber M, et al. Pivotal results of the Medtronic Vascular Talent Thoracic Stent Graft System: the VALOR trial. *J Vasc Surg* 2008; 48: 546–554.
27. Greenberg RK, O'Neill S, Walker E, et al. Endovascular repair of thoracic aortic lesions with the Zenith TX1 and TX2 thoracic grafts: intermediate-term results. *J Vasc Surg* 2005; 41: 589–596.
28. Matsumura JS, Cambria RP, Dake MD, et al. International controlled clinical trial of thoracic endovascular aneurysm repair with the Zenith TX2 endovascular graft: 1-year results. *J Vasc Surg* 2008; 47: 247–257.
29. Bavaria JE, Appoo JJ, Makaroun MS, et al. Endovascular stent grafting versus open surgical repair of descending thoracic aortic aneurysms in low-risk patients: a multicenter comparative trial. *J Thorac Cardiovasc Surg* 2007; 133: 369–377.
30. Makaroun MS, Dillavou ED, Wheatley GH, Cambria RA. Five-year results of endovascular treatment with the Gore TAG device compared to open repair of thoracic aortic aneurysms. *J Vasc Surg* 2008; 47: 912–918.
31. Appoo JJ, Moser WG, Fairman RM, et al. Thoracic aortic stent grafting: improving results with newer generation investigational devices. *J Thorac Cardiovasc Surg* 2006; 131: 1087–1094.
32. Kieffer E, Chiche L, Cormier E, Guegan H. Recurrent spinal cord ischemia after endovascular stent grafting for chronic traumatic aneurysm of the aortic isthmus. *J Vasc Surg* 2007; 45: 831–833.
33. Cho JS, Rhee RY, Makaroun MS. Delayed paraplegia 10 months after endovascular repair of thoracic aortic aneurysm. *J Vasc Surg* 2008; 47: 625–628.
34. Sheehan MK, Ouriel K, Greenberg R, et al. Are type II endoleaks after endovascular aneurysm repair endograft dependent? *J Vasc Surg* 2006; 43: 657–661.
35. Cho JS, Dillavou ED, Rhee RY, Makaroun MS. Late abdominal aortic aneurysm enlargement after endovascular repair with the Excluder device. *J Vasc Surg* 2004; 39: 1236–1241.
36. Haider SE, Najjar SF, Cho JS, et al. Sac behavior after aneurysm treatment with the Gore Excluder low-permeability aortic endoprosthesis: 12-month comparison to the original Excluder device. *J Vasc Surg* 2006; 44: 694–700.
37. Leurs LJ, Bell R, Degrieck Y, et al. Endovascular treatment of thoracic aortic diseases: combined experience from the EUROSTAR and United Kingdom Thoracic Endograft registries. *J Vasc Surg* 2004; 40: 670–679.
38. Resch T, Koul B, Dias N, et al. Changes in aneurysm morphology and stent-graft configuration after endovascular repair of aneurysms of the descending thoracic aorta. *J Thorac Cardiovasc Surg* 2001; 122: 47–52.
39. Akin I, Kische S, Ince H, Nienaber CA. Indication, timing and results of endovascular treatment of type B dissection. *Eur J Vasc Endovasc Surg* 2009; 37: 289–296.
40. Szeto WY, McGarvey M, Pochettino A, et al. Results of a new surgical paradigm: endovascular repair for acute complicated type B aortic dissection. *Ann Thorac Surg* 2008; 86: 87–94.
41. Eggebrecht H, Nienaber CA, Neuhauser M, et al. Endovascular stent-graft placement in aortic dissection: a meta-analysis. *Eur Heart J* 2006; 27: 489–498.
42. Schoder M, Czerny M, Cejna M, et al. Endovascular repair of acute type B aortic dissection: long-term follow-up of true and false lumen diameter changes. *Ann Thorac Surg* 2007; 83: 1059–1066.
43. Hsu R, Ho Y, Chen RJ, et al. Outcome of medical and surgical treatment in patients with acute type B aortic dissection. *Ann Thorac Surg* 2005; 79: 790–795.
44. Dake MD, Kato N, Mitchell RS, et al. Endovascular stent-graft placement for the treatment of acute aortic dissection. *N Engl J Med* 1999; 340: 1546–1552.
45. Nienaber CA, Fattori R, Lund G, et al. Nonsurgical reconstruction of thoracic aortic dissection by stent-graft placement. *N Engl J Med* 1999; 340: 1539–1545.
46. Tsai TT, Fattori R, Trimarchi S, et al. Long-term survival in patients presenting with type B acute aortic dissection: insights from the International Registry of Acute Aortic Dissection. *Circulation* 2006; 114: 2226–2231.
47. Nienaber CA, Kische S, Zeller T, et al. Provisional extension to induce complete attachment after stent-graft placement in type B aortic dissection: the PETTICOAT concept. *J Endovasc Ther* 2006; 13: 738–746.

Mount Sinai Journal of Medicine 77:270–278, 2010

Minimally Invasive Vein Surgery: Latest Options for Vein Disease

Steven Elias, MD, FACPh

Mount Sinai School of Medicine, New York, NY

OUTLINE

Diagnosis
Treatments
  Saphenous Incompetence
  Varicose Veins
  Perforator Incompetence
  Deep Disease
Conclusion

ABSTRACT

The goal of treatment for venous disease is to decrease ambulatory venous hypertension. Various strategies are employed. These can be divided into exogenous and endogenous treatments. Exogenous methods concern those employed from the outside of the limb, such as compression and elevation. Endogenous modalities treat from inside the limb the underlying venous pathology due to venous valvular dysfunction or venous obstruction. Traditional endogenous procedures include stripping, ligation, and phlebectomy. All these procedures require incisions, anesthesia, and perhaps hospitalization, and involve significant discomfort. Newer minimally invasive vein surgery procedures now exist. These are all same-day, outpatient procedures, usually involving local anesthesia. Most can be performed percutaneously without incisions. Patients ambulate the day of the procedure. Morbidity is less than 1%. This article summarizes the concept of minimally invasive vein surgery and summarizes new technologies to manage all forms of venous disease. ***Mt Sinai J Med 77:270–278, 2010.*** *© 2010 Mount Sinai School of Medicine*

Address Correspondence to:

Steven Elias
Mount Sinai School of Medicine
New York, NY
Email: steven.elias@mountsinai.org

Key Words: endogenous treatment, exogenous treatment, minimally invasive vein surgery, venous disease.

Venous disease is thought to have been reported in texts more than 3000 years old. Leg ulcers were first described by Hippocrates, who noted an association with varicose veins and ulceration.[1] He recommended those with leg ulcers should avoid standing. This concept of relieving venous hypertension persists, and today relief can be achieved with more rapid healing, minimal morbidity, and less recurrence.

The prevalence of venous disease far exceeds that of arterial disease, with 4× to 5× more cases of venous disease than arterial disease. Venous insufficiency may present as a spectrum of clinical disease, from spider veins to long-standing venous stasis ulcers. Approximately 20% of men and 40% of women may experience some form of vein disease. The clinical-etiology-anatomy-pathophysiology (CEAP) classification based on manifestations of chronic venous disease was developed in 1994 and revised in 2004.[2,3] The CEAP system, which ranges from asymptomatic (C0) to active ulcer (C6), is the most widely used descriptive system for venous disease and provides a systematic guide for its investigation (Table 1). The Edinburgh Vein Study found the prevalence of chronic venous insufficiency (CVI) was 9.4% in men and 6.6% in women.[4] The San Diego Population Study found venous trophic changes in 6.2% of the studied population.[5] McLafferty found a CEAP C4 represented 8% of the population, whereas CEAP C5 or C6 was found in 1.5%.[6] Venous stasis ulceration has been estimated to represent 58% to 70% of lower-extremity ulcers.[7]

The prevalence of venous ulcers has ranged in various studies from 0.1% to 0.63% of the general population.[8] In the United States, this potentially represents up to 1.8 million people with ulcerative venous disease. The prevalence of venous ulcers increases with increasing age. This likely reflects the progression of venous disease leading

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20186

© 2010 Mount Sinai School of Medicine

Table 1. *Clinical-Etiology-Anatomy-Pathophysiology Classification for Chronic Venous Disease.*

Classification	Description
C0	Asymptomatic; no visible or palpable signs of disease
C1	Telangiactasias, or reticular or spider veins (<0.3 mm)
C2	Varicose veins
C3	Edema of the legs
C4	Changes in skin and subcutaneous tissue
4A	Pigmentation or eczema
4B	Lipodermatosclerosis or atrophie blanche
C5	Healed, closed venous ulcer
C6	Active, open venous ulcer

to the development of lower-extremity ulcers. As the population ages, the overall number of patients suffering from advanced vein disease is likely to increase.

DIAGNOSIS

Duplex ultrasound is the diagnostic modality of choice to confirm presence of venous disease. Venous insufficiency is indicated by reversal of flow or valve closure time of >0.5 second in the superficial system and >1.0 second in the deep system. In the perforator system, a reversal of flow time of >0.5 second with a vein diameter of 3.0 mm indicates incompetence, and regardless of flow dynamics, perforator veins with a diameter >4.0 mm should be considered incompetent.[9] Incompetence does not necessarily equate to a perforator being pathologic. This issue needs further investigation considering flow velocities and flow volume. Mixed disease may be present and can affect the therapies appropriate for ulcer management. Patients with mild arterial disease (an ankle-brachial index [ABI] >0.8) in the presence of venous ulcers should be able to tolerate adequate compression and definitive venous procedures without arterial compromise to the limb. In those with an ABI of 0.5 to 0.8, the patient may be treated with reduced compression, that is, 15 to 25 mm Hg and evaluation of the arterial system.[10]

At times, computed tomographic (CT) venography, magnetic resonance (MR) venography or ascending venography, and intravascular ultrasound are utilized to evaluate the pelvic venous system for outflow disease secondary to iliac-vein stenosis or obstruction. Patients seek treatment for symptomatic disease, cosmetic improvement, or management of complications of disease. At times patient goals may overlap to include improvement of both appearance and symptoms. The complications of venous disease that prompt treatment are superficial or deep phlebitis, thrombosis, variceal bleeding, skin changes, or ulceration.

Minimally invasive vein surgery (MIVS) options can be divided into the venous systems they treat: superficial, perforator, or deep. Superficial techniques address visible varicose veins or their underlying cause (great or small saphenous incompetence). Perforator management directly targets refluxing perforator veins. Deep procedures treat venous obstruction/stenosis or incompetence.

TREATMENTS

Saphenous Incompetence

Surgery for the superficial venous system of the lower extremity has evolved significantly in recent years. For the great and small saphenous veins, stripping and ligation was the most popular method for most of the 20th century. The standard procedure involves the exposure of the distal saphenous vein at the level of the knee or ankle and passage of a stripper to the level of the sapheno-femoral junction. The great saphenous vein is then ligated, stripped, and removed through the skin incision. Similar techniques can also be employed for the small saphenous vein if incompetence is present. The procedure is done on a same-day basis, with a patient recovery time of 2 to 3 weeks. Saphenous vein stripping and ligation is a well-established technique with proven benefits in treatment of venous disease.

Less-invasive alternatives to vein ligation and stripping developed in the year 2000 and have been employed widely since. Endovenous ablation using either radiofrequency (RF) (Figure 1) or laser energy (Figure 2) has been performed at increasing rates and now supersedes conventional vein stripping and ligation. The successful treatment of superficial incompetence by endovenous ablation has been shown to be comparable to results from surgery in several randomized controlled trials.[11,12] The reported advantages of these techniques are decreased pain and bruising at the operative site, no incisions, and decreased recovery time for the patient. Complication rates are very low in endovenous ablation. The most common are minor and include superficial thrombophlebitis, pain, cellulitis, and hematoma. Deep vein thrombosis is a rare complication that can occur when thrombus induced by venous ablation extends into the common femoral vein.[13] We find this complication can be avoided by performing ablation an

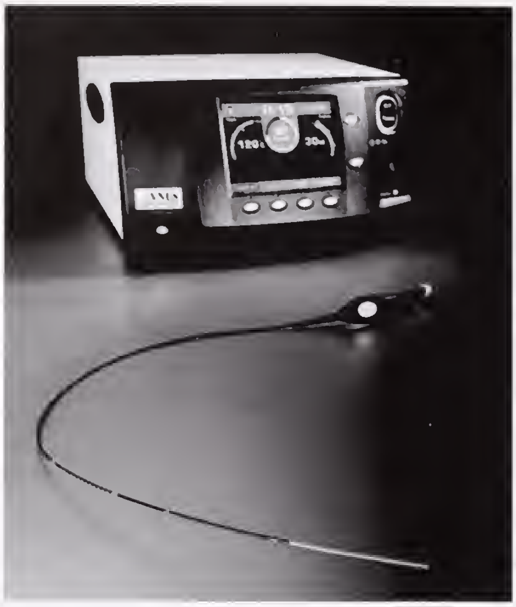


Fig 1. Radiofrequency catheter.

Fig 2. Laser fiber.

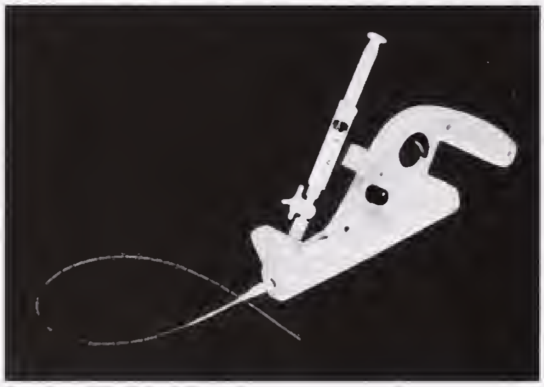


Fig 3. ClariVein catheter.

adequate distance (2 cm) from the sapheno-femoral junction or sapheno-popliteal junction. Compared with stripping, endovenous ablation also reduces recurrence rates (neovascularization) in the groin.

> *For most of the 20th century, the most popular method for treating saphenous incompetence was stripping and ligation. The preferred method now is endovenous ablation using either radiofrequency or laser energy. As shown in several randomized controlled trials, this method produces results comparable with those of surgery. The reported advantages are decreased pain and bruising at the operative site, no incisions, and decreased recovery time. Complication rates are very low.*

Important factors in the success of endovenous ablation are accurate access and delivery of adequate energy for contraction of the vein wall and ablation of the lumen. The 2 energy-delivery devices available use either radiofrequency (VNUS Medical Technologies) or laser energy. The newest device to treat saphenous incompetence is ClariVein (Vascular Insights) (Figure 3). This utilizes mechanical and chemical occlusion of the veins by a rotating intraluminal wire and liquid sclerosant. This author has completed the initial first-in-man clinical trial.[14] It appears that initial 6-month results are comparable to laser or radiofrequency ablation.

Varicose Veins

The removal of visible varicose veins has traditionally been the mainstay of management. Many times by ablating saphenous venous incompetence with previously mentioned methods, the connected varicose veins reduce because the underlying venous hypertension has decreased. However, at times it is appropriate to remove the varicosities to decrease the abnormal venous burden or to improve cosmesis.

Ambulatory or microphlebectomy has been utilized with great success recently. This technique uses small, punctate incisions and tiny hooks that allow the veins to be pulled above the skin and removed. This procedure can be time consuming and tedious, especially with extensive, large varicose veins.

Another MIVS technique called TriVex (Smith & Nephew) is a mechanical device to remove visible veins. This has been available since 2000. One can think of the technique as "liposuction of the veins." There are 2 components: a transilluminator/irrigator and a resector (Figure 4). The transilluminator allows the operator to "see" the veins and the resector sucks the veins into a rotating blade where they are morcellated and removed. Operative time is short (15 minutes) and the number of incisions per patient is small (average 2.5).[15] Prospective randomized trial results compare well with ambulatory phlebectomy.[16]

In summary, combining MIVS for saphenous treatment and visible varicosities allows the practitioner to significantly improve patient symptoms and cosmetic appearance. This is accomplished with minimal incisions and minimal time for patient recovery. All of these procedures are performed with local anesthesia, with sedation in some cases. Patients can return to full activity in 1 to 2 days.

Combining MIVS for saphenous treatment and visible varicosities allows the practitioner to significantly improve patient symptoms and cosmetic appearance. This is accomplished with minimal incisions and minimal time for patient recovery. All of these procedures are performed with local anesthesia, with sedation in some cases. Patients can return to full activity in 1 to 2 days.

Perforator Incompetence

Incompetent perforator veins (IPVs) may significantly contribute to lower-extremity venous hypertension. Perforator incompetence is thought to arise from several possible causes, such as post-thrombophlebitic changes or longstanding venous incompetence as a result of deep vein thrombosis. The presence of IPVs in venous ulcer disease may exacerbate symptoms, prevent ulcer healing, or predispose to recurrence. The treatment of IPVs has traditionally employed open exposure and subfascial ligation, also known as the Linton procedure.[17] Open perforator ligation showed a high wound complication rate of 20% to 50% as well as high recurrence rates of up to 22% at 50 months' follow-up.[18] Postoperative pain was usually significant, requiring narcotic analgesics, and often involved an in-hospital stay of 3 to 5 days.

The traditional method for the treatment of incompetent perforator veins (IPVs) is open exposure and subfascial ligation. This method involves a high wound complication rate (20% to 50%), high recurrence rates (up to 22% at 50 months), significant postoperative pain, and in-hospital stays of 3 to 5 days.

In an attempt to minimize the drawbacks of IPV ligation, endoscopic techniques were developed. Subfascial endoscopic perforator surgery (SEPS) involves the endoscopic access and identification of IPVs and their ligation.[19–21] One of the advantages of this technique over the open procedure are that the skin incisions are remote from the site of ulcer, thus avoiding the damaged and potentially infected ulcerated skin. The decreased incisional size also reduces postoperative pain. The reported wound complication rates of 6% to 7% are much lower than for the open procedure. SEPS improves on the technique of perforator ligation and is done as an outpatient procedure. The disadvantages of the procedure include need for regional or general anesthesia and incidence of nerve injury. It can be technically difficult to perform and the accessibility to perforators near the malleolus can be limited because of angulation of the endoscopic instruments in the small subfascial space.

The advantages of subfascial endoscopic perforator surgery (SEPS) for IPVs are skin incisions remote from the site of ulcer, decreased incisional size and postoperative pain, and performance as an outpatient procedure. The disadvantages are the need for regional or general anesthesia and incidence of nerve injury.

Advances in endovenous treatment of veins have allowed for the development of percutaneous ablation of perforators (PAPs). This recently coined term[22] was described conceptually as early as 1974, when liquid sclerosant was used in ablation of incompetent perforating veins, initially with promising results.[23–25] PAPs may be performed by several methods, although all have the same objectives: (1) ultrasound-guided percutaneous intraluminal access; (2) application of ablative technique within leading to vessel occlusion and vein wall contraction, (3) local anesthesia with or without sedation, (4) outpatient setting, with no incisions; and (5) simple retreatment if needed.

The absence of an incision and wound should yield a near-zero wound infection and complication rate. No dissection of tissue is done, as opposed to SEPS or open surgery. Postoperative pain is minimal and readily managed by oral, non-narcotic analgesics. Recurrence of incompetence or development of new incompetent perforators has been well described, and multiple interventions may be required.[26,27] If they are, previous performance of PAPs should not make reintervention more difficult, although reintervention may be more difficult in SEPS or open surgery. PAPs is easily repeatable, with minimal morbidity.

In contrast to SEPS and open surgery, percutaneous ablation of perforators (PAPs) involves no incision, near-zero wound infection and complication rates, and minimal postoperative pain, which can be readily managed by oral, non-narcotic analgesics. Moreover, PAPs is easily repeatable, with minimal morbidity.

In SEPS, anatomic considerations may limit veins amenable to ligation. Veins near the malleolus are technically difficult to access, even for those experienced in SEPS. In PAPs, intervention is limited only by proper ultrasound visualization of the perforating vein. Perimalleolar and even inframalleolar perforators can be accessed and treated. PAPs may also be better accepted by the physician and patient because of the minimal pain and morbidity associated with this technique.

Theoretical complications of PAPs include skin injury, nerve injury, and deep-vessel injury. The use of thermal energy may lead to collateral damage of surrounding structures, but this author has not seen these complications occur. Both PAPs and SEPS have a steep learning curve. The vessels to be treated are generally in the range of 3 to 5 mm and can be difficult to access.

Techniques of Percutaneous Ablation of Perforators

All modalities of PAPs adhere to a basic methodology: (1) ultrasound-guided intraluminal access, (2) delivery of ablative measure (thermal or chemical), (3) confirmation of initial treatment success, and (4) follow-up assessment of treatment success.

All modalities of PAPs adhere to a basic methodology: (1) ultrasound-guided intraluminal access, (2) delivery of ablative measure (thermal or chemical), (3) confirmation of initial treatment success, and (4) follow-up assessment of treatment success.

Ultrasound-guided access may be obtained by needle, angiocatheter, or the radiofrequency stylet (RFS) catheter (VNUS Medical Technologies). The RFS device is currently the only proprietary catheter specifically available for perforator treatment (Figure 5). Laser ablation is currently considered off-label use. For all devices, entry into the IPV is

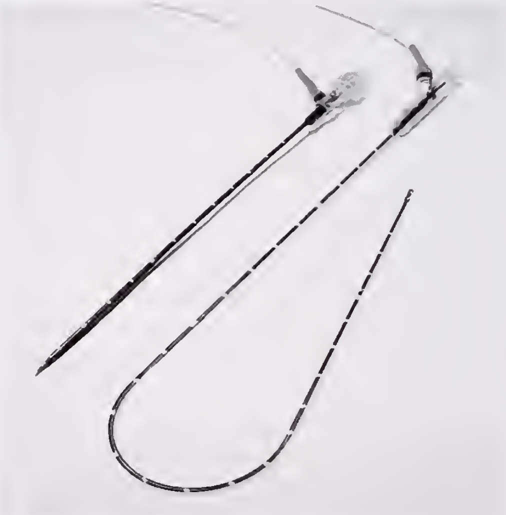


Fig 5. RFS catheter.

DOI:10.1002/MSJ

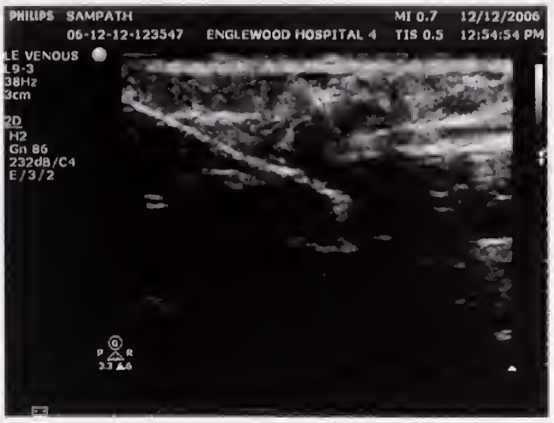

PHILIPS SAMPATH
06-12-12-123547 ENGLEWOOD HOSPITAL 4
MI 0.7 12/12/2006
TIS 0.5 12:54:54 PM
LE VENOUS
L9-3
38Hz
3cm
2D
H2
Gn 86
232dB/C4
E/3/2

Fig 6. Catheter in perforator lumen.

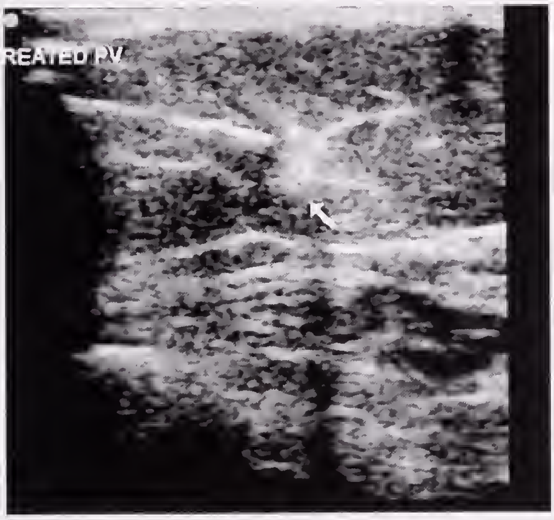


Fig 7. Changes in perforator after treament.

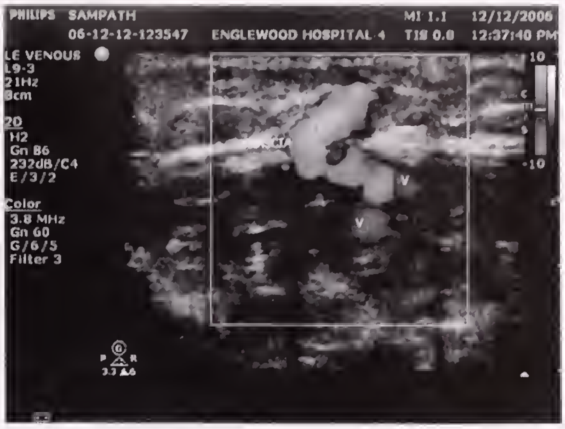

PHILIPS SAMPATH
06-12-12-123547 ENGLEWOOD HOSPITAL 4
MI 1.1 12/12/2006
TIS 0.8 12:37:40 PM
LE VENOUS
L9-3
21Hz
3cm
2D
H2
Gn 86
232dB/C4
E/3/2
Color
3.8 MHz
Gn 60
G/6/5
Filter 3
FASCIA
V
V

Fig 8. Pre treatment perforator with flow.

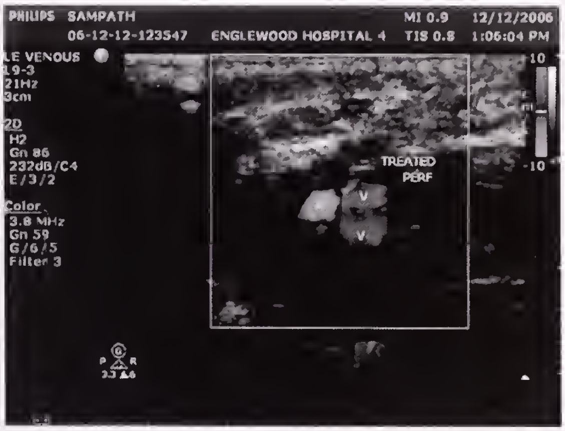

PHILIPS SAMPATH
06-12-12-123547 ENGLEWOOD HOSPITAL 4
MI 0.9 12/12/2006
TIS 0.8 1:06:04 PM
LE VENOUS
L9-3
21Hz
3cm
2D
H2
Gn 86
232dB/C4
E/3/2
Color
3.8 MHz
Gn 59
G/6/5
Filter 3
TREATED PERF
V
V

Fig 9. Post treatment perforator with no flow.

confirmed visually and by aspiration of blood from the lumen (Figure 6). In addition, the RFS catheter has the ability to measure impedance in ohms and is used as a further confirmatory parameter; measurement of <400 ohms indicates intraluminal position. This measurement significantly enhances successful cannulation. Thermal energy is then delivered to ablate the target vessel (Figure 7). The RF catheter heats the vein to 85°C for 4 minutes in each treated segment. In the laser system, approximately 120 to 150 joules should be delivered to the vein. A sufficient length (2 to 3 mm) should be treated to assure that the segment just below the fascia but remote from the deep system is ablated. Post-PAPs ultrasound confirms ablation (Figure 8 and Figure 9). Compression dressing is then applied over the treated areas. When sclerotherapy is used, access is obtained as described. The sclerosant used can be in either liquid form or, more recently, foam. Foam sclerotherapy has been advocated as more effective. Most studies use sodium tetradecyl sulfate (STS) or sodium morrhuate in liquid form, injecting volumes of 0.5 to 1 mL sclerosant. Care is taken not to inject into the accompanying artery. After infusion, compression is applied with direct pressure over the treated IPV, followed by dressing with wraps or stockings.

Clinical Results of Percutaneous Ablation of Perforators

Sclerotherapy in PAPs has been employed since 1963, although early attempts did not use ultrasound guidance.[28] Thibault and Lewis used ultrasound-guided sclerotherapy with an 85% IPV closure rate at 6 months.[24] Subsequent studies with a range of sclerosants have found immediate treatment success rates ranging from 90% to 98% and IPV closure rates of up to 75% at 20 months.[29]

Using ultrasound-guided sclerotherapy, Thibault and Lewis demonstrated an 85% IPV closure rate at 6 months. Subsequent studies with a range of sclerosants have found immediate treatment success rates ranging from 90% to 98% and IPV closure rates of up to 75% at 20 months.

Limited data exist regarding RF ablation of perforators. Whiteley and colleagues performed 770 IPV ablations in 506 limbs with 79% occlusion at 1 year.[30] Following this report, the RFS designed specifically for IPV ablation was approved by the Food and Drug Administration in 2004. Early data from Chang and colleagues' use of early RFS showed a 100% procedural success rate. Follow-up in 38 perforators at 6 months showed 87% reflux free, and at 1 year 91% were reflux free.[31] Our own as-yet-unpublished series involved 20 IPVs treated for CEAP class C4 to C6 venous disease. The 1-month occlusion rate was 100%, and the 4-month rate was 90%. The 4-month ulcer healing rate was 90%. Peden reports treatment in 69 IPVs with better than 90% success at 6-month follow-up.[32] The early data for laser treatment show comparable results. A report by Proebstle and Herdemann of IPV treatment using laser energy involved 67 IPVs in 60 limbs.[33] The majority of patients had CEAP class C2 disease, and 1 patient had CEAP class C6. Procedural occlusion rate was 100%. Three-month follow-up included only 16 IPVs but did show continued occlusion in these veins. Kabnick reported a 4-month closure rate of 85% in 25 IPVs treated with laser energy.[34] Our own results treating 50 IPVs in CEAP classes C4 to C6 with laser energy showed 1-month closure rate of 90%.[32] Murphy compared 100 IPVs treated with RF and 100 IPVs treated with laser and showed at 6-month follow-up a 90% closure rate for RF and 100% closure for laser.[35] The ablation of IPVs using thermal energy is a relatively new procedure, and long-term data are not yet available, but the early (6-month to 1-year) results of PAPs have been promising. Chemical ablation has been performed for longer and has, generally, had good results also. The principles of endovenous treatment for great and small saphenous veins apply to PAPs as well: Good technique and adequate delivery of thermal energy provide optimal treatment.

Deep Disease

A significant number of venous ulcer patients have deep insufficiency, sclerosis, or obstruction. In the recent past, the entity of proximal suprainguinal disease was not appreciated. Neglen has brought this pathology to treating physicians' attention.[36] Outflow stenosis or obstruction in the iliac venous system causes downstream venous hypertension. Catheter-based technology is now available to identify and treat outflow problems. These techniques include ascending venography, MR venography, and intravascular ultrasound. Treatment consists of balloon angioplasty and stent placement with good long-term success. Percutaneous prosthetic valves may soon be available for implantation in patients with deep valvular insufficiency.[37] These valves will also allow the reconstruction of an incompetent deep venous system in a percutaneous manner. The treatment of the deep system pursues the common goal of reduction of venous pressures (Figure 10). The MIVS management of deep disease is the last frontier of modern care of venous disease. It is beyond the scope of this article, but percutaneous MIVS methods now exist for prompt treatment of DVT. These modalities ultimately should decrease the incidence of post-thrombotic syndrome.

CONCLUSION

Since the beginning of the 21st century, there now exist MIVS methods to treat all 3 venous systems.

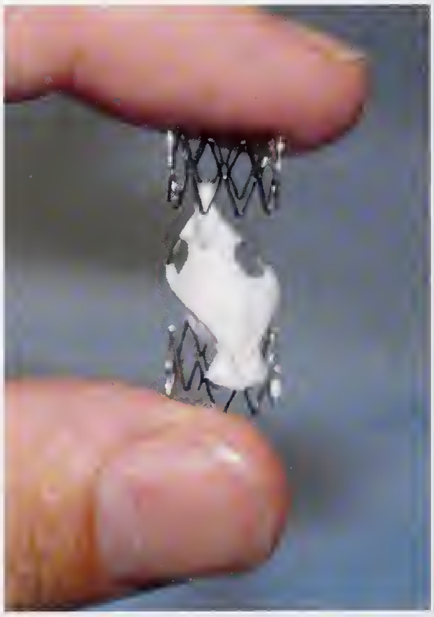


Fig 10. Percutaneous "Sail" valve.

The great majority of these techniques achieve good if not better results than traditional open methods employed for almost a century. All of the described procedures utilize either no incisions or minimal (3 mm) incisions. Almost all are outpatient, same-day procedures employing local anesthesia or intravenous sedation. Patients can return to normal activities within 1 to 2 days. It is important that physicians and patients understand the advances in venous disease care with methods of MIVS.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Hippocrates. De ulceribus and de carnibus. In: Adams F, ed. *The Genuine Works of Hippocrates*. London, UK: Sydenham Society; 1849; 325–336.
2. Beebe HG, Bergan JJ, Bergqvist D, et al. Classification and grading of chronic venous disease in the lower limbs: a consensus statement. *Eur J Vasc Endovasc Surg* 1996; 12: 487–491.
3. Eklof B, Rutherford RB, Bergan JJ, et al. Revision of the CEAP classification for chronic venous disorders: consensus statement. *J Vasc Surg* 2004; 40: 1248–1252.
4. Evans CJ, Fowkes FG, Ruckley CV, Lee AJ. Prevalence of varicose veins and chronic venous insufficiency in men and women in the general population: Edinburgh Vein Study. *J Epidemiol Community Health* 1999; 53: 149–153.
5. Criqui MH, Jamosmos M, Fronek A, et al. Chronic venous disease in an ethnically diverse population: the San Diego Population Study. *Am J Epidemiol* 2003; 158: 448–456.
6. McLafferty RB, Lohr JM, Caprini JA, et al. Results of the national pilot screening program for venous disease by the American Venous Forum. *J Vasc Surg* 2007; 45: 142–148.
7. Barwell JR, Davies CE, Deacon J, et al. Comparison of surgery and compression with compression alone in chronic venous ulceration (ESCHAR study): randomised controlled trial. *Lancet* 2004; 363: 1854–1859.
8. Coleridge Smith PD. *Leg Ulcers*. 3rd ed. London, UK: Oxford University Press; 2005.
9. Elias SM, Frasier KL. Minimally invasive vein surgery: its role in the treatment of venous stasis ulceration. *Am J Surg* 2004;188(suppl 1A): 26–30.
10. Marston W. Evaluation and treatment of leg ulcers associated with chronic venous insufficiency. *Clin Plast Surg* 2007; 34: 717–730.
11. Rautio T, Ohinmaa A, Perala J, et al. Endovenous obliteration versus conventional stripping operation in the treatment of primary varicose veins: a randomized controlled trial with comparison of the costs. *J Vasc Surg* 2002; 35: 958–965.
12. Darwood RJ, Theivacumar N, Dellagrammaticas D, et al. Randomized clinical trial comparing endovenous laser ablation with surgery for the treatment of primary great saphenous varicose veins. *Br J Surg* 2008; 95: 294–301.
13. Puggioni A, Kalra M, Carmo M, et al. Endovenous laser therapy and radiofrequency ablation of the great saphenous vein: analysis of early efficacy and complications. *J Vasc Surg* 2005; 42: 488–493.
14. Elias S. Mechanical-chemical endovenous ablation (ClariVein): a new tumescentless technique. Poster at: American College of Phlebology 23rd Annual Congress; November 5–8, 2009; Palm Desert, CA.
15. Elias S. Transilluminated powered phlebectomy. In: Bergan JJ, ed. *The Vein Book*. Philadelphia, PA: Elsevier; 2007.
16. Aremu MA, Mahendran B, Butcher W, et al. Prospective randomized controlled trial: conventional versus powered phlebectomy. *J Vasc Surg* 2004;39: 88–94.
17. Linton RR. The operative treatment of varicose veins and ulcers, based upon a classification of these lesions. *Ann Surg* 1938; 107: 582–593.
18. Pierik EG, van Urk H, Hop WC, Wittens CH. Endoscopic versus open subfascial division of incompetent perforating veins in the treatment of venous leg ulceration: a randomized trial. *J Vasc Surg* 1997; 26: 1049–1054.
19. Gloviczki P, Bergan JJ, Rhodes JM, et al. Mid-term results of endoscopic perforator vein interruption for chronic venous insufficiency: lessons learned from the North American subfascial endoscopic perforator surgery registry. The North American Study Group. *J Vasc Surg* 1999; 29: 489–502.
20. Ciostek P, Myrcha P, Noszczyk W. Ten years experience with subfascial endoscopic perforator vein surgery. *Ann Vasc Surg* 2002; 16: 480–487.
21. Ting AC, Cheng SW, Ho P, et al. Clinical outcomes and changes in venous hemodynamics after subfascial endoscopic perforating vein surgery. *Surg Endosc* 2003; 17: 1314–1318.
22. Elias SM. Will SEPS be around in the next 5 years? Paper presented at: International Vein Conference; April 2005; Miami, FL.
23. Hobbs JT. Surgery and sclerotherapy in the treatment of varicose veins: a random trial. *Arch Surg* 1974; 109: 793–796.
24. Thibault PK, Lewis WA. Recurrent varicose veins. Part 2: Injection of incompetent perforating veins using ultrasound guidance. *J Dermatol Surg Oncol* 1992; 18: 895–900.
25. Guex JJ. Ultrasound guided sclerotherapy (USGS) for perforating veins (PV). *Hawaii Med J* 2000; 5: 261–262.
26. Roka F, Binder M, Bohler-Sommeregger K. Mid-term recurrence rate of incompetent perforating veins after combined superficial vein surgery and subfascial endoscopic perforating vein surgery. *J Vasc Surg* 2006; 44: 359–363.
27. Van Rij AM, Hill G, Gray C, et al. A prospective study of the fate of venous leg perforators after varicose vein surgery. *J Vasc Surg* 2005; 42: 1156–1162.
28. Fegan WG. Continuous compression technique of injecting varicose veins. *Lancet* 1963; 2: 109–112.
29. Masuda EM, Kessler DM, Lurie F, et al. The effect of ultrasound-guided sclerotherapy of incompetent perforator veins on venous clinical severity and disability scores. *J Vasc Surg* 2006; 43: 551–556.

30. Whiteley MS, Holdstock JM, Price BA, et al. Radiofrequency ablation of refluxing great saphenous systems, Giacomini veins, and incompetent perforating veins using VNUS closure and TRLOP technique. *Phlebology* 2003; 18: 52.
31. Chang DW, Levy D, Hayashi RM, et al. Percutaneous catheter treatment of perforating vein incompetence using ultrasound guided radiofrequency ablation. Paper presented at: Society for Clinical Vascular Surgery Annual Meeting; March 9–13, 2005;; Coral Gables, FL.
32. Elias S, Peden E. Ultrasound-guided percutaneous ablation for the treatment of perforating vein incompetence. *Vascular* 2007; 15: 281–289.
33. Proebstle TM, Herdemann S. Early results and feasibility of incompetent perforator vein ablation by endovenous laser treatment. *Dermatol Surg* 2007; 33: 162–168.
34. Kabnick L. Perforator vein treatment. Paper presented at: InVein Meeting; June 2006; Uncasville, CT.
35. Murphy R. Comparison of radiofrequency and laser for perforator treatment. Paper presented at: American College of Phlebology 20th Annual Congress; November 9–12, 2006; Ponte Vedra, FL.
36. Neglen P. Chronic deep venous obstruction: definition, prevalence, diagnosis, management. *Phlebology* 2008; 23: 149–157.
37. Teebken OE, Puschmann C, Aper T, et al. Tissue-engineered bioprosthetic venous valve: a long-term study in sheep. *Eur J Vasc Endovasc Surg* 2003; 25: 305–312.

Mount Sinai Journal of Medicine 77:279–285, 2010

Intracranial Aneurysms: Endovascular Treatment

Henry Moyle, MD, PhD, and Aman B. Patel, MD

Mount Sinai School of Medicine, New York, NY

OUTLINE

- Saccular Intracranial Aneurysm Coil Embolization
- Endovascular Treatment for Fusiform/Dissecting Aneurysms
- Conclusion

ABSTRACT

The incidence of subarachnoid hemorrhage is estimated at 5 to 10 per 100,000 per year. In patients who survive the initial hemorrhage, the repeat hemorrhage rate is 15% to 20% in the first 2 weeks after presentation and is associated with devastating clinical outcomes even graver than the initial rupture. The current options for aneurysm treatment are surgical clipping and, since the mid-1990s, neuroendovascular coil embolization. The former was at one time the gold standard of care, but consistent with the trend in modern medicine toward less-invasive procedures, the latter has steadily gained prominence. Although there is still controversy as to its long-term durability and safety, it is now the preferred procedure. This article describes the procedure's antecedents, rationale, and essential components. ***Mt Sinai J Med 77:279–285, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: fusiform-dissecting aneurysm, neuroendovascular coil embolization, saccular intracranial aneurysm, subarachnoid hemorrhage, surgical clipping.

Address Correspondence to:

Aman B. Patel
Mount Sinai School of Medicine
New York, NY
Email: aman.patel@mountsinai.org

The incidence of subarachnoid hemorrhage (SAH) is estimated at 5 to 10 per 100,000 per year.[1] Saccular intracranial aneurysm rupture accounts for 80% of spontaneous SAH. SAH is associated with devastating clinical outcomes; about 50% of patients will die within a month of rupture, and only about one-third of survivors will have a good clinical outcome.[2,3] In patients who survive the initial hemorrhage, the repeat hemorrhage rate is 15% to 20% in the first 2 weeks after presentation and is associated with devastating clinical outcomes even graver than the initial rupture. Early treatment aimed at preventing subsequent rupture and repeat hemorrhage is the current standard of care. Currently there are 2 options for aneurysm treatment: surgical clipping and, since the mid-1990s, neuroendovascular coil embolization. Prior to neuroendovascular coil embolization of aneurysms, surgical clipping was the gold standard of treatment. Surgical clipping involves a craniotomy and microdissection of the aneurysm to allow clip placement across the neck of the aneurysm, resulting in exclusion of the aneurysm sac from the normal parent vessel. Generally, as trends in modern medicine have moved toward less-invasive procedures, aneurysm treatment has followed with the development of neuroendovascular approaches. These methods take their roots from early work on peripheral embolization of vascular malformations and facial tumors utilizing embolic metallic coils. Transarterial microcatheters are used to gain access to the aneurysm and then deliver embolic materials into the aneurysm sac without compromising the parent vessel (Figure 1). Although there is still controversy as to the long-term durability and safety of endovascular coil embolization, its role continues to expand and dominate as the preferred procedure for aneurysm treatment, with an estimated 200,000 performed as of 2003.[4]

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20183

© 2010 Mount Sinai School of Medicine

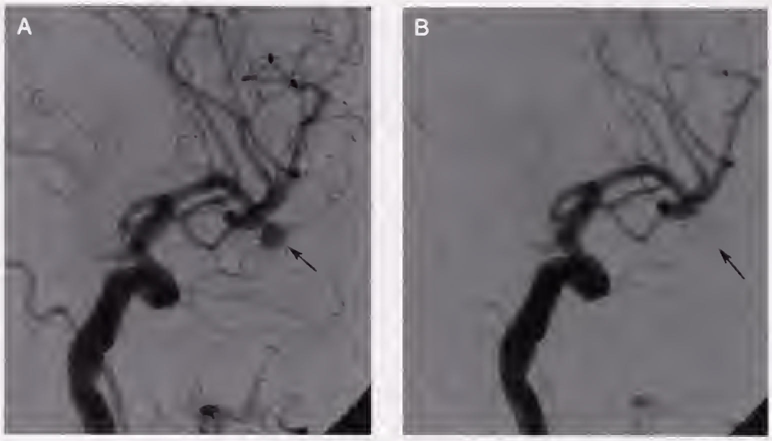


Fig 1. Standard cerebral angiogram with arrows demonstrating a left middle cerebral artery aneurysm (A) filling with contrast prior to treatment and (B) after platinum coil embolization.

There are currently 2 options for aneurysm treatment: surgical clipping and, since the mid-1990s, neuroendovascular coil embolization. At one time, the former was the gold standard of treatment. Now, the latter is the preferred procedure, with an estimated 200,000 performed as of 2003.

SACCULAR INTRACRANIAL ANEURYSM COIL EMBOLIZATION

Although early neuroendovascular work started in the 1970s, complete or near-complete aneurysm occlusion lagged until the early 1990s as the characteristics of coil delivery, stiffness, and shape memory improved. During the early 1990s, several papers were published describing successes with this technique[5,6] culminating with the work of Guido Guglielmi,[7,8] who described the use of the first detachable coils in the treatment of intracranial aneurysms. The electrolytic detachment system of these coils allowed for control of coil deployment into the aneurysm. This system also minimized the risk of coil migration and dislodgement into distal arterial territories.[7,8] Guglielmi's development led to the release of the first US Food and Drug Administration–approved detachable coil, the Guglielmi Detachable Coil (Boston Scientific Neurovascular, Fremont, CA), in 1994. Detachable coils have now become the current standard embolic material for endovascular treatment of intracranial aneurysms, with multiple coil systems currently on the market.

In 2005, the International Subarachnoid Trial (ISAT) published its initial results comparing surgical clipping with coil embolization in a prospective randomized trial. ISAT reported a relative risk reduction for dependency or death of 23.9% with an absolute risk reduction of 7.4% in patients undergoing coiling. This study concluded that, in patients with ruptured intracranial aneurysms suitable for either treatment, a policy of endovascular coiling was substantially more likely to result in independent survival at 1 year compared with neurosurgical clipping.[9] More recent ISAT subgroup analysis also suggests that age at onset of aneurysm rupture may be important in determining superiority of coil embolization over surgical clipping.[10] The International Study on Unruptured Intracranial Aneurysms (ISUIA) has also suggested coil embolization is a valid alternative to surgery, especially in patients older than 50 years, with larger aneurysms (>10 mm), and with aneurysms of the posterior circulation.[11] Despite growing evidence that coil embolization is a viable alternative to surgical clipping, shortcomings of bare platinum coils still exist. The ISUIA study revealed that by using coil embolization with bare platinum detachable coils, only 55% of aneurysms were obliterated completely,

In 2005, the International Subarachnoid Trial reported that, compared with surgical clipping, coiling reduced the relative risk of dependency or death by 23.9% and the absolute risk by 7.4%.

DOI:10.1002/MSJ

24% were obliterated partially, and 18% could not be treated.

> *Despite growing evidence that coil embolization is a viable alternative to surgical clipping, shortcomings of bare platinum coils still exist. The International Study on Unruptured Intracranial Aneurysms (ISUIA) revealed that by using coil embolization with bare platinum detachable coils, only 55% of aneurysms were obliterated completely, 24% were obliterated partially, and 18% could not be treated.*

Incomplete aneurysm obliteration, coil compaction, and aneurysmal recanalization, the latter of which may occur in up to 40% of cases, are concerns.[12–14] These concerns may be particularly true in cases of large aneurysms or aneurysms with broad necks (dome/neck ratio <1.5 or neck >4 mm).[12–14] Because of the clinically devastating results of aneurysmal hemorrhage, long-term durability of aneurysm occlusion and low recurrent bleeding rates reflect efficacy and safety of coil embolization, respectively. These issues were addressed in a more recent publication based on 1810 patient-years follow-up demonstrating a 2.2% periprocedural morbidity rate and a 0.2% rate of recurrent bleeding.[15] The Cerebral Aneurysm Rerupture After Treatment study reports a 2.2% rerupture rate in the first year after coil embolization, followed by 0.2% and 0% in the second and third years, respectively.[16] Both studies strongly suggest rebleeding is related to the initial degree of aneurysm occlusion, but this remains controversial, as other investigators believe the angiographic stability of neck remnants over time is more important than the remnant itself.[15]

The issues of incomplete obliteration and recanalization have been addressed by advances in coil design. In particular, more complex-shape coil designs have entered the market with the hopes of obtaining higher aneurysm packing densities.[17,18] It is felt this may reduce recanalization, particularly in larger aneurysms. Another design modification in coils has been bioactive or coated coils. These devices include coils coated with polyglycolic acid (PGA) or combined polyglycolic/polylactic acid (PGLA), nylon/Dacron (Dupont, Wilmington, DE)/PGLAfibers, and hydrogel coating. The concept behind these coated coils has been to promote neointimal formation with subsequent fibrosis protecting the neck of the aneurysm[19] or to increase the percentage of the volume occluded. Although these coils have been available since 2002, randomized controlled studies are still underway.[20] The Hydrocoil Endovascular Aneurysm Occlusion and Packing Study trial (HELPS) is the largest randomized trial using coated coils to date. Follow-up data is not yet available, but procedural adverse events are similar to many case series already published for coil embolization.[21]

> *More complex-shape coil designs have entered the market with the hopes of obtaining higher aneurysm packing densities, which may reduce recanalization, particularly in larger aneurysms.*

Broad-necked aneurysms (>4 mm or dome/neck ratio <2) present a challenge to endovascular coil embolization techniques. When coiling such aneurysms, prolapse of coils through the wide neck into the parent vessel is a technical limitation. Failure of neointimal fibrosis across the wide neck can result in delayed recanalization. Again, with improved technology, these issues have been addressed by several techniques, including balloon or stent-assisted coil embolization and, more recently, liquid embolic agents. Temporary inflation of a nondetachable balloon across the aneurysm neck has been used in securing placement of the initial coils within wide-necked aneurysms, as a rescue method in cases of coil prolapse, and to allow for optimal packing of the aneurysm fundus and neck region (Figure 2).[22] Although thromboembolic complications were thought to be higher using this adjunctive technique, literature meta-analysis did not demonstrate higher incidence of thromboembolic events or iatrogenic aneurysm rupture.[23] This technique is more attractive in the setting of an acute ruptured aneurysm, as it has the advantage of not requiring long-term or premedication with antiplatelet therapy, unlike stent-assisted coil embolization. The development of self-expanding intracranial stents, the Neuroform stent (Boston Scientific/Target Therapeutics, Fremont, CA) and the Enterprise Vascular Reconstruction Device and Delivery System (Codman, Raynham, MA), has allowed for coil embolization of wide-necked aneurysms (>4 mm) and aneurysms with dome/neck ratios <2.[24–26] The stent is thought to provide a buttress to prevent coil prolapse into the parent vessel (Figure 3). The

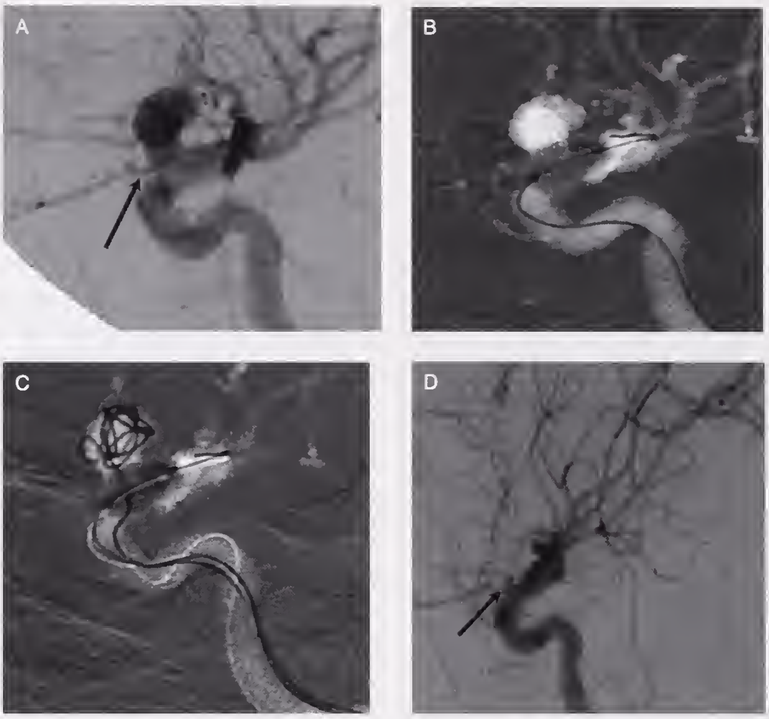


Fig 2. Standard cerebral angiogram demonstrating a wide-necked carotid-ophthalmic aneurysm. The arrow (A) points out the origin of the ophthalmic artery incorporated into the base of the aneurysm. (B) A HyperForm balloon (HyperForm, Micro Therapeutics, Irvine, CA) was inflated across the neck, remodeling the neck to preserve the ophthalmic artery. (C) Coil embolization was then performed with the balloon inflated to protect the origin of the ophthalmic artery. (D) Postembolization, the ophthalmic artery is preserved, as demonstrated by the arrow.

stent may also provide flow diversion away from the aneurysm sac[27] and potentially provide scaffolding for orifice endothelialization.[28] Results of the largest series to date, a 142-patient multicenter registry, indicate 2.8% morbidity and 2% mortality with stent-assisted coil embolization. This technique is preferred in unruptured aneurysms, as patients are premedicated and maintained on antiplatelet therapy to prevent thromboembolic complications.

The Onyx liquid embolic system (Micro Therapeutics, Inc., Irvine, CA) has been used in the treatment of intracranial arteriovenous malformations and aneurysms since 1999.[29,30] Onyx is a nonadhesive biocompatible liquid embolic agent constituted by ethylene-vinyl alcohol copolymer with dimethyl sulfoxide solvent added to microlyzed tantalum powder, which is a high-molecular-weight metal. The combination makes Onyx radiopaque. Onyx can be delivered through a microcatheter directly into the aneurysm, where the liquid quickly transforms into a solid polymer cast. Leakage into the parent vessel is prevented by combining delivery with balloon remodeling of the aneurysm neck. In 2004, Molyneux *et al.* reported the results of a multicenter evaluation using Onyx for the embolization of intracranial aneurysms. At 12 months, there was more than 90% total or subtotal occlusion.[31] The properties of Onyx counter 3 major difficulties in the treatment of wide-necked aneurysms with platinum coils: (1) it produces total or near-total filling of the aneurysm sac; (2) it provides aneurysm neck and parent vessel reconstruction (by the balloon remodeling technique discussed earlier); and (3) it induces a stronger aneurysm neck reaction of neoendothelialization as compared with regular platinum coils.[29] In a 69-patient study, morbidity and mortality rates are similar to other endovascular techniques.[32]

ENDOVASCULAR TREATMENT FOR FUSIFORM/DISSECTING ANEURYSMS

Fusiform/dissecting aneurysms differ from saccular aneurysms in that the latter arise from a focal defect in the parent vessel wall, whereas the former arise

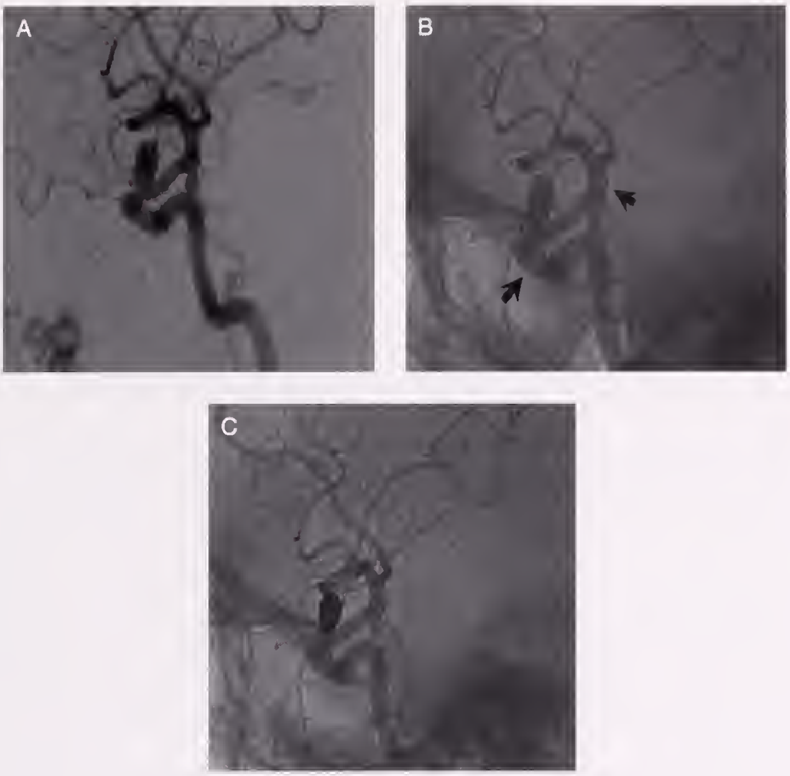


Fig 3. Standard cerebral angiogram demonstrating (A) a wide-necked internal carotid artery aneurysm. (B) An Enterprise stent (Enterprise Vascular Reconstruction Device; Cordis Neurovascular, Inc., Miami Lakes, FL) was placed across the neck, with arrows demonstrating the radiopaque markers built into the ends of the stent. (C) Coil embolization was subsequently performed.

from larger, more diffuse segmental defects in the parent vessel wall, usually involving >25% of the vessel circumference. Endosaccular embolic approaches do not address this large dysplastic parent vessel. The self-expanding intracranial stents previously discussed have been used with some success in cases of fusiform/dissecting aneurysms.[33] A new class of stent called the Pipeline embolization device (Chestnut Medical Technologies, Inc., Menlo Park, CA) has been engineered to act as a stand-alone device for the endovascular reconstruction of a segmentally diseased parent vessel. This device has 30% to 35% metal surface area coverage when deployed, as opposed to 6.5% to 9% coverage with the Neuroform or Enterprise self-expanding stents. This device represents a shift from endosaccular embolization to flow diversion and endothelialization of the dysplastic segment in the parent vessel. A recent study of 53 patients treated with this device demonstrated 95% total occlusion at 12 months; however, only 18 patients were followed to the 12-month endpoint.[34] Other studies note similar encouraging results.[35–38] The concept of flow diversion stents may represent a new and completely different paradigm for endovascular intracranial aneurysm treatment.

CONCLUSION

The field of neuroendovascular treatment of intracranial aneurysms has seen a rapid progression in technological advancements over the last 20 years. For ruptured aneurysms that are suited for surgical or endovascular treatment, there is class I evidence to support decreased dependency or death with endovascular treatment when compared with surgical treatment. Technological advancements have also allowed endovascular treatment for more complex aneurysms once thought only amenable to surgery. Although certain aneurysms will still require surgical intervention, preferred treatment when possible appears to be less-invasive endovascular approaches.

DISCLOSURES

Potential conflict of interest: Aman B. Patel has served as a Codman Neurovascular consultant.

REFERENCES

1. Broderick JP, Brott TG, Duldner JE, et al. Volume of intracerebral hemorrhage: a powerful and easy-to-use predictor of 30-day mortality. *Stroke* 1993; 24: 987–993.
2. Hop JW, Rinkel GJ, Algra A, van Gijn J. Case-fatality rates and functional outcome after subarachnoid hemorrhage: a systematic review. *Stroke* 1997; 28: 660–664.
3. Sarti C, Tuomilehto J, Salomaa V, et al. Epidemiology of subarachnoid hemorrhage in Finland from 1983 to 1985. *Stroke* 1991; 22: 848–853.
4. Friedman JA, Nichols DA, Meyer FB, et al. Guglielmi detachable coil treatment of ruptured saccular cerebral aneurysms: retrospective review of a 10-year single-center experience. *AJNR Am J Neuroradiol* 2003; 24: 526–533.
5. Dowd CF, Halbach VV, Higashida RT, et al. Endovascular coil embolization of unusual posterior inferior cerebellar artery aneurysms. *Neurosurgery* 1990; 27: 954–961.
6. Hilal SK, Solomon RA. Endovascular treatment of aneurysms with coils. *J Neurosurg* 1992; 76: 337–339.
7. Guglielmi G, Vinuela F, Dion J, Duckwiler G. Electrothrombosis of saccular aneurysms via endovascular approach. Part 2: preliminary clinical experience. *J Neurosurg* 1991; 75: 8–14.
8. Guglielmi G, Vinuela F, Sepetka I, Macellari V. Electrothrombosis of saccular aneurysms via endovascular approach. Part 1: electrochemical basis, technique, and experimental results. *J Neurosurg* 1991; 75: 1–7.
9. Molyneux AJ, Kerr RS, Yu LM, et al. International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised comparison of effects on survival, dependency, seizures, rebleeding, subgroups, and aneurysm occlusion. *Lancet* 2005; 366: 809–817.
10. Ryttlefors M, Enblad P, Kerr RS, Molyneux AJ. International subarachnoid aneurysm trial of neurosurgical clipping versus endovascular coiling: subgroup analysis of 278 elderly patients. *Stroke* 2008; 39: 2720–2726.
11. Unruptured intracranial aneurysms–risk of rupture and risks of surgical intervention. International Study of Unruptured Intracranial Aneurysms Investigators [published correction appears in *N Engl J Med* 1999; 340: 744]. *N Engl J Med* 1998; 339: 1725–1733.
12. Cognard C, Weill A, Spelle L, et al. Long-term angiographic follow-up of 169 intracranial berry aneurysms occluded with detachable coils. *Radiology* 1999; 212: 348–356.
13. Sluzewski M, van Rooij WJ, Rinkel GJ, Wijnalda D. Endovascular treatment of ruptured intracranial aneurysms with detachable coils: long-term clinical and serial angiographic results. *Radiology* 2003; 227: 720–724.
14. Vinuela F, Duckwiler G, Mawad M. Guglielmi detachable coil embolization of acute intracranial aneurysm: perioperative anatomical and clinical outcome in 403 patients. *J Neurosurg* 1997; 86: 475–482.
15. Holmin S, Krings T, Ozanne A, et al. Intradural saccular aneurysms treated by Guglielmi detachable bare coils at a single institution between 1993 and 2005: clinical long-term follow-up for a total of 1810 patient-years in relation to morphological treatment results. *Stroke* 2008; 39: 2288–2297.
16. Johnston SC, Dowd CF, Higashida RT, et al. Predictors of rehemorrhage after treatment of ruptured intracranial aneurysms: the Cerebral Aneurysm Rerupture After Treatment (CARAT) study. *Stroke* 2008; 39: 120–125.
17. Hirsch JA, Bendok BR, Paulsen RD, et al. Midterm clinical experience with a complex-shaped detachable platinum coil system for the treatment of cerebral aneurysms: Trufill DCS Orbit detachable coil system registry interim results. *J Vasc Interv Radiol* 2007; 18: 1487–1494.
18. Wakhloo AK, Gounis MJ, Sandhu JS, et al. Complex-shaped platinum coils for brain aneurysms: higher packing density, improved biomechanical stability, and midterm angiographic outcome. *AJNR Am J Neuroradiol* 2007; 28: 1395–1400.
19. Pierot L, Leclerc X, Bonafe A, Bracard S; for French Matrix Registry Investigators. Endovascular treatment of intracranial aneurysms using Matrix coils: short- and midterm results in ruptured and unruptured aneurysms. *Neurosurgery* 2008; 63: 850–858.
20. White PM, Raymond J. Endovascular coiling of cerebral aneurysms using "bioactive" or coated-coil technologies: a systematic review of the literature. *AJNR Am J Neuroradiol* 2009; 30: 219–226.
21. White PM, Lewis SC, Nahser H, et al. HELPS Trial Collaboration. HydroCoil Endovascular Aneurysm Occlusion and Packing Study (HELPS trial): procedural safety and operator-assessed efficacy results. *AJNR Am J Neuroradiol* 2008; 29: 217–223.
22. Aletich VA, Debrun GM, Misra M, et al. The remodeling technique of balloon-assisted Guglielmi detachable coil placement in wide-necked aneurysms: experience at the University of Illinois at Chicago. *J Neurosurg* 2000; 93: 388–396.
23. Shapiro M, Babb J, Becske T, Nelson PK. Safety and efficacy of adjunctive balloon remodeling during endovascular treatment of intracranial aneurysms: a literature review. *AJNR Am J Neuroradiol* 2008; 29: 1777–1781.
24. Lubicz B, Francois O, Levivier M, et al. Preliminary experience with the Enterprise stent for endovascular treatment of complex intracranial aneurysms: potential advantages and limiting characteristics. *Neurosurgery* 2008; 62: 1063–1070.
25. Mocco J, Snyder KV, Albuquerque FC, et al. Treatment of intracranial aneurysms with the Enterprise stent: a multicenter registry. *J Neurosurg* 2009; 110: 35–39.
26. Peluso JP, van Rooij WJ, Sluzewski M, Beute GN. A new self-expandable nitinol stent for the treatment of wide-neck aneurysms: initial clinical experience. *AJNR Am J Neuroradiol* 2008; 29: 1405–1408.
27. Kim M, Levy EI, Meng H, Hopkins LN. Quantification of hemodynamic changes induced by virtual placement of multiple stents across a wide-necked basilar trunk aneurysm. *Neurosurgery* 2007; 61: 1305–1313.
28. Biondi A, Janardhan V, Katz JM, et al. Neuroform stent-assisted coil embolization of wide-neck intracranial aneurysms: strategies in stent deployment and midterm follow-up. *Neurosurgery* 2007; 61: 460–469.
29. Murayama Y, Vinuela F, Tateshima S, et al. Endovascular treatment of experimental aneurysms by use of a combination of liquid embolic agents and protective devices. *AJNR Am J Neuroradiol* 2000; 21: 1726–1735.
30. Murayama Y, Vinuela F, Ulhoa A, et al. Nonadhesive liquid embolic agent for cerebral arteriovenous

malformations: preliminary histopathological studies in swine rete mirabile. *Neurosurgery* 1998; 43: 1164–1175.

31. Molyneux AJ, Cekirge S, Saatci I, Gal G. Cerebral Aneurysm Multicenter European Onyx (CAMEO) trial: results of a prospective observational study in 20 European centers. *AJNR Am J Neuroradiol* 2004; 25: 39–51.
32. Piske RL, Kanashiro LH, Paschoal E, et al. Evaluation of Onyx HD-500 embolic system in the treatment of 84 wide-neck intracranial aneurysms. *Neurosurgery* 2009; 64: E865–E875.
33. Wakhloo AK, Mandell J, Gounis MJ, et al. Stent-assisted reconstructive endovascular repair of cranial fusiform atherosclerotic and dissecting aneurysms: long-term clinical and angiographic follow-up. *Stroke* 2008; 39: 3288–3296.
34. Lylyk P, Miranda C, Ceratto R, et al. Curative endovascular reconstruction of cerebral aneurysms with the pipeline embolization device: the Buenos Aires experience. *Neurosurgery* 2009; 64: 632–643.
35. Fiorella D, Kelly ME, Albuquerque FC, Nelson PK. Curative reconstruction of a giant midbasilar trunk aneurysm with the pipeline embolization device. *Neurosurgery* 2009; 64: 212–217.
36. Fiorella D, Woo HH, Albuquerque FC, Nelson PK. Definitive reconstruction of circumferential, fusiform intracranial aneurysms with the pipeline embolization device. *Neurosurgery* 2008; 62: 1115–1121.
37. Hoit DA, Schirmer CM, Malek AM. Stent graft treatment of cerebrovascular wall defects: intermediate-term clinical and angiographic results. *Neurosurgery* 2008; 62(5 suppl 2): ONS380–ONS389.
38. Pumar JM, Lete I, Pardo MI, et al. LEO stent monotherapy for the endovascular reconstruction of fusiform aneurysms of the middle cerebral artery. *AJNR Am J Neuroradiol* 2008; 29: 1775–1776.

Deep Vein Thrombosis: Endovascular Management

Robert A. Lookstein, MD, and Christopher F. Giordano, MD

Division of Interventional Radiology, Mount Sinai School of Medicine, New York, NY

OUTLINE

NATURAL HISTORY OF DEEP VEIN THROMBOSIS AND POST-THROMBOTIC SYNDROME
EVIDENCE OF THE BENEFITS OF THROMBUS REMOVAL
COMPLICATIONS, ESTABLISHED INDICATIONS, AND CONTRAINDICATIONS OF CATHETER-DIRECTED THROMBOLYSIS
  Complications
  Indications
  Contraindications
HISTORICAL PERSPECTIVE OF TREATMENT STRATEGIES OF THROMBUS REMOVAL
  Systemic Thrombolysis
  Catheter-Directed Thrombolysis
  Open Surgical Thrombectomy
  Percutaneous Mechanical Thrombectomy
  Pharmacomechanical Thrombectomy
  Venous Angioplasty and Stenting
FUTURE PERSPECTIVES
CONCLUSION

ABSTRACT

Deep vein thrombosis is a condition that affects hundreds of thousands of patients each year. The major complications include pulmonary embolus with the potential for loss of life and post-thrombotic syndrome with the potential for loss of function of the limb. Extensive clinical research over the last 40 years has improved the techniques to remove the thrombus from the affected limb and reduce the likelihood of developing post-thrombotic syndrome. These treatments have included intravenous systemic thrombolysis, catheter-directed thrombolysis, surgical thrombectomy, and most recently pharmacomechanical thrombectomy to rapidly fragment, lyse, and remove the thrombus from the affected limb. This last technique may finally transform the treatment of acute deep vein thrombus from strictly conservative medical therapy to a minimally invasive procedure that can remove the thrombus to improve the quality of life of millions of individuals suffering from the symptoms of deep vein thrombosis. ***Mt Sinai J Med 77:286–295, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: deep vein thrombosis, endovascular, thrombectomy, thrombolysis.

Address Correspondence to:

Robert A. Lookstein
Mount Sinai School of Medicine
New York, NY
Email: robert.lookstein@mssm.edu

Acute deep vein thrombosis (DVT) is defined by the Society of Interventional Radiology as partial or complete thrombosis of a deep vein in which symptoms have been present for 14 days or fewer or when imaging studies indicate that venous thrombosis has occurred within the past 14 days.[1] Lower-extremity DVT is a grave condition that can result in major disability or death due to pulmonary embolism, post-thrombotic syndrome (PTS), paradoxical embolization, or phlegmasia cerulea dolens leading to limb loss.[2] Each year approximately 2 million patients in the United States will develop DVT and 200,000 will subsequently die of pulmonary embolism, with the incidence of pulmonary embolism–related mortality increasing with age, regardless of sex and race.[3] In addition to the early complications of acute DVT, late-morbidity

Lower-extremity deep vein thrombosis (DVT) is a grave condition that can result in major disability or death due to pulmonary embolism, post-thrombotic syndrome, paradoxical embolization, or phlegmasia cerulea dolens leading to limb loss.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20184

© 2010 Mount Sinai School of Medicine

complications may develop from recurrent thrombosis and PTS.[4]

DVT traditionally has been treated with anticoagulation (heparin) and lower-extremity elastic compression stockings since the 1950s. This medical therapy remains an important component of the contemporary DVT treatment paradigm. Anticoagulant therapy prevents recurrent ipsilateral DVT, a major risk factor for pulmonary embolism and PTS.[5] Heparin and the low-molecular-weight heparins prevent thrombus propagation, which allow the body's own fibrinolytic system to help remove the thrombus. However, heparin by itself does not remove an existing thrombus from the deep venous system, which led to the use of intravenous thrombolytics in the 1960s and then catheter-directed thrombolysis (CDT) in the 1980s in an effort to remove the thrombus from the deep venous circulation. Surgical thrombectomy and then percutaneous mechanical thrombectomy were explored in the 1990s with varying rates of success and morbidity. The combination of CDT and mechanical thrombectomy (pharmacomechanical thrombectomy) is the contemporary technique of thrombus removal. Balloon angioplasty and endovascular stent placement are also being increasingly used in the management of acute and chronic DVT. The natural history of DVT and PTS, and the indications and contraindications of thrombus removal, are important to understand to help guide patient care.

NATURAL HISTORY OF DEEP VEIN THROMBOSIS AND POST-THROMBOTIC SYNDROME

Physicians may fail to distinguish the difference in the pathophysiology of primary versus post-thrombotic venous insufficiency and therefore do not appreciate the rationale for thrombus removal. The acute phase of DVT presents with leg edema, pain, and difficulty ambulating, which arises from venous hypertension caused by outflow obstruction.[6] These symptoms typically subside over days to weeks because of collateral flow.[6] Patients may experience outflow obstruction indefinitely. A primary goal of thrombus removal is relief of outflow obstruction, which does not occur entirely through anticoagulation and lower-extremity compression stockings because thrombus regression occurs in only 50% of patients.[7,8] The most common clinical manifestation of this chronic obstruction is PTS.

PTS is a long-term consequence of DVT and commonly includes chronic limb edema, heaviness, pain, lifestyle-limiting venous claudication, and stasis dermatitis.[9] Approximately 20% to 50% of patients develop PTS within 1 to 2 years of symptomatic DVT, and 5% to 10% of patients develop severe PTS, which includes venous ulcers.[10] The pathophysiology of PTS is ambulatory venous hypertension, which is defined as an elevated venous pressure during exercise.[11] Venous valvular incompetence and luminal obstruction contribute to ambulatory venous hypertension.[12]

Studies have consistently shown that patients with chronic venous obstruction have the most severe morbidity from PTS[11,13] and that multisegment venous involvement and iliofemoral obstruction specifically result in the most profound morbidity.[14,15] The natural history of iliofemoral DVT carries a worse prognosis compared with isolated femoropopliteal DVT. This is because spontaneous lysis and recanalization of the femoropopliteal venous segment and development of collateral vessels often enhance venous outflow and limit the severity of PTS. In comparison, iliofemoral DVT rarely recanalizes via spontaneous lysis and recanalization, and persistent venous outflow obstruction often leads to major increases in ambulatory venous pressures.[16] Recurrent DVT was observed to occur twice as frequently in patients with iliofemoral DVT than in those with infrainguinal DVT.[17] In a long-term study of patients with iliofemoral DVT, 44% experienced significant venous claudication during a median follow-up of 5 years.[18]

PTS has profound socioeconomic impact via increased medical costs and the indirect costs of work disability. A 1997 Swedish study demonstrated that the average overall direct medical cost of treatment of late complications of DVT was US $4659, or approximately 75% of the cost of treatment of the primary DVT.[19] Venous ulcer development was the most costly complication.[19] The direct cost of treatment of chronic venous disease is thought to exceed $300 million yearly in the United States, and it is estimated that 2 million workdays are lost annually as a result of leg ulcers.[20] These studies confirm the severe clinical disability and high socioeconomic costs associated with PTS, and its profound impact as a major health care problem for which a better preventive strategy is needed.

EVIDENCE OF THE BENEFITS OF THROMBUS REMOVAL

There are small single-center series that suggest that thrombus removal or early thrombus resolution after the onset of acute DVT is associated with improved

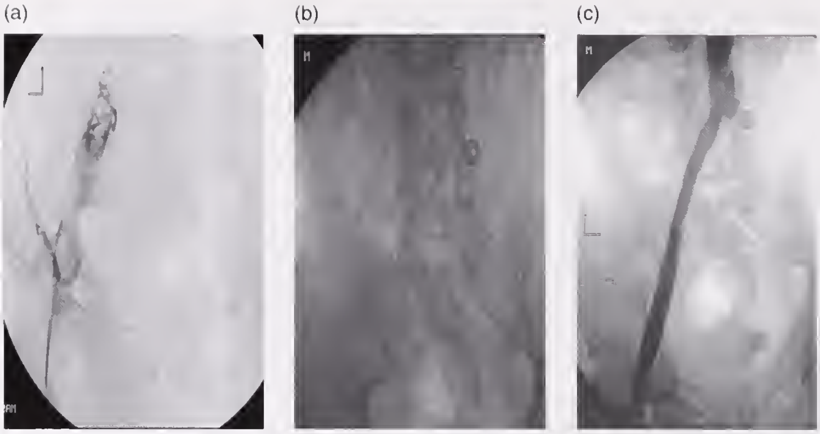


Fig 1. A 32-year-old male with lower-extremity swelling for 10 days. (A) The initial image is a conventional venogram performed from a left popliteal approach with the patient in the prone position. The entire left external and common iliac veins are thrombosed. Following pharmacomechanical thrombectomy, stenosis (not shown) was identified in the left common iliac vein secondary to May-Thurner syndrome. May-Thurner syndrome is thickening and intimal hyperplasia of the left common iliac vein caused by the arterial pressure of the crossing right common iliac artery. (B) Placement of an endovascular stent in the left common iliac vein stenosis. (C) Conventional venogram demonstrates widely patent veins with brisk, straight flow into the inferior vena cava.

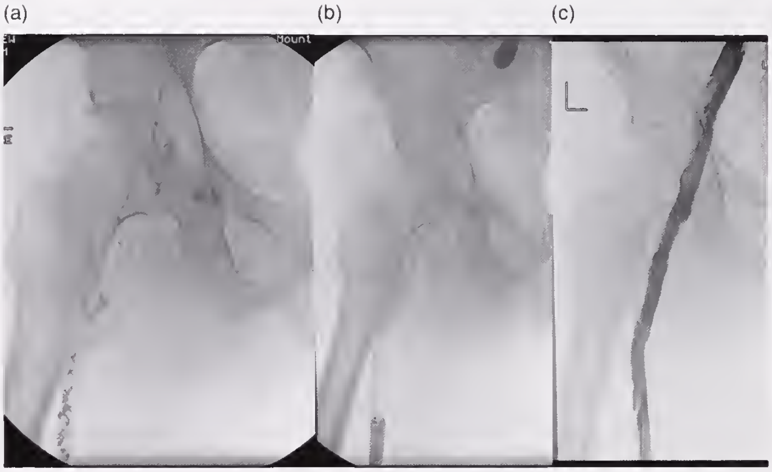


Fig 2. A 30-year-old male with 1 week of painful, swollen left leg. (A) Conventional venogram demonstrates complete thrombosis of the left femoral and common femoral vein. (B) Following placement of the Bacchus Vascular Trellis isolated thrombolysis system. (C) Following treatment, the venous segment is widely patent without significant residual thrombus. This procedure was completed in a single session without the need for overnight infusion of a thrombolytic drug.

outcomes.[12] Experimental observations of acute DVT in canine models demonstrated that thrombolysis preserves endothelial function and valve competence immediately and at 4 weeks after therapy compared with placebo. In addition, there was less residual thrombus in veins treated with a plasminogen activator, thereby preserving the vein's structural integrity.[21,22]

A study of the natural history of acute DVT in patients treated with anticoagulation alone found that valvular incompetence developed in the femoral veins with persistent iliofemoral vein obstruction, even when femoral veins had no thrombus involvement.[23] The combination of valvular incompetence and venous obstruction was associated with the most severe post-thrombotic

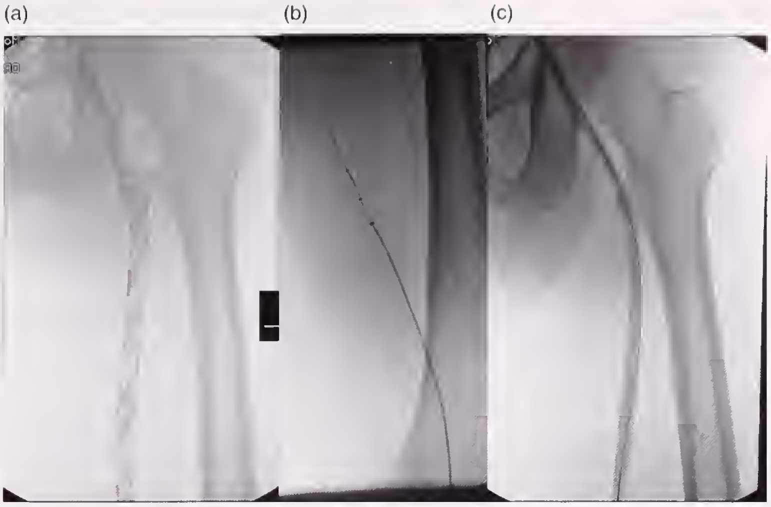


Fig 3. A 46-year-old male with a painful swollen right leg for three days. (A) The initial image is a conventional venogram performed from a right popliteal approach with the patient in the prone position. The entire right femoral and common femoral veins are thrombosed. (B) After treatment with pharmacomechanical thrombectomy with the Possis AngioJet using powerpulse spray technique. (C) The vein is widely patent without any residual thrombus. This procedure was completed on the same day without the need for overnight infusion of a thrombolytic drug.

morbidity.[24] Spontaneous clot lysis naturally restored venous patency, and if spontaneous lysis occurred early ($\leq$90 days), valve function was frequently preserved.[25]

> *A study of the natural history of acute DVT in patients treated with anticoagulation alone found that valvular incompetence developed in the femoral veins with persistent iliofemoral vein obstruction, even when femoral veins had no thrombus involvement.*

Persistent thrombus at termination of anticoagulation is associated with recurrent DVT.[12] Recurrent DVT significantly increases the risk of post-thrombotic morbidity, implying that removal of thrombus will decrease the risk of PTS.[12]

Long-term efficacy of thrombus removal in patients with acute iliofemoral DVT was evaluated by a prospective randomized trial of iliofemoral venous surgical thrombectomy with creation of a popliteal arteriovenous fistula and anticoagulation versus anticoagulation alone.[26] The creation of an arteriovenous fistula is believed to augment blood flow across the thrombogenic luminal venous surface.[26] Follow-up at 6 months, 5 years, and 10 years demonstrated clear benefit in patients randomized to early venous thrombectomy.[26,27] These benefits included improved patency of the iliofemoral venous system, lower venous pressures, less edema, and fewer post-thrombotic symptoms.[26,27] A 10-year follow-up showed consistent benefit of thrombectomy, but the number of patients surviving to 10 years was small, thereby reducing the statistical power of these observations.[26]

> *In a prospective randomized trial of iliofemoral venous surgical thrombectomy with creation of a popliteal arteriovenous fistula and anticoagulation versus anticoagulation alone, follow-up at 6 months, 5 years, and 10 years demonstrated clear benefit in patients randomized to early venous thrombectomy.*

The National Venous Thrombolysis Registry (NVTR) demonstrated impressive results in terms of success of thrombus dissolution as well as improved quality of life. This study included 287 patients treated with catheter-directed urokinase, and the mean follow-up was 1 year. Complete dissolution of thrombus was noted in 31% of cases and partial (50% to 99%) thrombus dissolution in an additional 52%, and primary patency at 1 year was 60%.[28] Patency was greater in iliofemoral segments than in femoropopliteal segments and in patients in whom complete dissolution of thrombus occurred during the initial hospitalization.[6] An important observation was made in the subgroup of patients with acute, first-time iliofemoral DVT who had initially successful thrombolysis: 96% of the veins in these patients remained patent at 1 year.[12] Valvular competence was preserved in 72% of patients with complete thrombolysis.[28]

A follow-up study of the NVTR demonstrated better functioning and quality of life in patients with iliofemoral DVT treated by CDT than in those treated with anticoagulants alone.[29] The quality-of-life results were directly related to the initial success of thrombolysis.[12] These patients who had a successful lytic outcome reported a significantly improved health utilities index, increased physical functioning, less stigmata of chronic venous disease, less health distress, and fewer overall post-thrombotic symptoms.[12] Patients in whom CDT failed had similar results to patients treated with anticoagulation alone.[30]

COMPLICATIONS, ESTABLISHED INDICATIONS, AND CONTRAINDICATIONS OF CATHETER-DIRECTED THROMBOLYSIS

Complications

Major bleeding is the most common major complication of CDT and occurs in approximately 8% of patients undergoing treatment.[2] The range of major bleeding rates (0% to 24%) is relatively wide, apparently due to differences in the respective patient populations, small sample size in the single-center studies, and variability in reporting criteria.[2] Intracranial bleeding, retroperitoneal bleeding, pulmonary embolism, and death represent the most clinically significant complications of CDT. However, analysis of the published literature indicates that each of these complications is rare (1% to 2%).[2]

Indications

The currently established indications for pharmacological CDT are phlegmasia cerulea dolens, acute/subacute inferior vena cava thrombosis, acute iliofemoral DVT, acute femoropopliteal DVT, and subacute/chronic iliofemoral DVT.[2] Phlegmasia cerulea dolens and acute/subacute inferior vena cava thrombosis are two indications of CDT performed even in those patients with moderate bleeding risks and shorter life expectancy due to the severe, acute consequences.[2] Phlegmasia refers to the distinctive clinical picture in which DVT causes massive swelling of the entire extremity.[31] This disease requires emergent treatment because it can lead to arterial insufficiency, compartment syndrome, and/or venous gangrene and ultimately limb amputation.[31] The goal of CDT in these cases is limb salvage.[2]

The currently established indications for pharmacological catheter-directed thrombolysis (CDT) are phlegmasia cerulea dolens, acute/subacute inferior vena cava thrombosis, acute iliofemoral DVT, acute femoropopliteal DVT, and subacute/chronic iliofemoral DVT.

Acute/subacute inferior vena cava thrombosis is another indication of CDT.[2] Patients typically present with moderate to severe pelvic congestion and limb symptoms, hepatic veno-occlusive disease (Budd-Chiari syndrome) or acute renal failure.[2] The presence of an inferior vena cava thrombus can also preclude the placement of an inferior vena cava filter for pulmonary emboli prophylaxis.[2] Surgical thrombectomy is reserved for patients with phlegmasia cerulea dolens or acute/subacute inferior vena cava thrombosis who are at high risk of bleeding.[2]

Acute iliofemoral DVT in ambulatory patients with low bleeding risk and long life expectancy is a third indication of CDT. Patients with acute iliofemoral DVT are usually highly symptomatic and are high-risk for recurrent DVT, PTS, and late disability.[17,18,32] CDT can remove venous thrombus, provide immediate symptom relief, and facilitate treatment of underlying venous stenoses with balloon angioplasty and/or endovascular stent placement (Fig. 1).[33] If patients have a moderate or high risk of bleeding, surgical thrombectomy may be considered.[2]

Acute femoropopliteal DVT in highly symptomatic ambulatory patients with low bleeding risk and long life expectancy is a fourth indication for CDT based on its potential to prevent or minimize

PTS.[2] The threshold for utilizing CDT in the treatment of acute femoropopliteal DVT should probably be higher than that for acute iliofemoral DVT. This is because the ability of CDT to remove venous thrombus and provide immediate symptom relief in patients with acute femoropopliteal DVT is similar to that observed in patients with acute iliofemoral DVT, but midterm venous patency rates have been lower in the femoropopliteal DVT patients.[34] Patients with femoropopliteal DVT may experience symptom resolution and have few or no manifestations of PTS when treated with anticoagulation and compression stockings alone.[18] The best candidates in this subgroup are those patients with progression of DVT despite anticoagulant therapy and those with severe symptoms, long life expectancy, and good performance status.[2]

Subacute and chronic iliofemoral DVT in patients with moderate to severe pelvic and/or limb symptomatology and low bleeding risk is the fifth, somewhat controversial, indication of CDT.[2] These patients suffer from the worst forms of PTS due to the combination of valvular reflux and persistent venous obstruction, which occurs because the iliac vein rarely recanalizes.[35,36] CDT is not expected to result in a normal limb in these patients, because the vein has usually sustained some degree of permanent damage. The primary goals of therapy are improvement in presenting symptoms, reduction of venous disability, healing of existing venous ulcers, and removal of the superimposed acute thrombus in patients with acute-on-chronic DVT.[2] Although there is reported success using CDT to treat patients with isolated chronic femoropopliteal DVT, these patients constitute a comparatively poor anatomic subgroup for endovascular treatment and the Society of Interventional Radiology encourages frequent review of the most current clinical treatment results.[2,37]

Contraindications

In general, CDT is contraindicated in any patient with a hemorrhagic disorder, an anatomical lesion that is prone to bleeding, or an absolute contraindication to anticoagulant therapy.[28] Some specific absolute contraindications include active bleeding, disseminated intravascular coagulation, recent cerebrovascular event (including transient ischemic attack), and neurosurgery or intracranial trauma within 3 months.[2] Some strong relative contraindications include recent cardiopulmonary resuscitation, organ biopsy, uncontrolled hypertension, recent major gastrointestinal bleeding within 3 months, massive pulmonary emboli, and known right-to-left cardiac or pulmonary shunt.[2]

HISTORICAL PERSPECTIVE OF TREATMENT STRATEGIES OF THROMBUS REMOVAL

Systemic Thrombolysis

The treatment of acute DVT with thrombolytic therapy to remove the thrombus was initially attempted by peripheral intravenous administration.[12] The systemic administration of thrombolytic agents, which is effective in small arteries such as the coronaries, is ineffective for treatment of DVT, likely because of inefficient diffusion of the thrombolytics into the matrix of large venous thrombi under conditions of low flow.[37–39] Thirteen studies have been reported comparing anticoagulant therapy with systemic thrombolytic therapy for acute DVT.[12] The diagnosis in these studies was established with ascending phlebography, which was repeated after treatment to assess outcome.[12] Pooled analysis indicates that only 4% of patients treated with anticoagulation alone had significant or complete lysis and an additional 14% had partial lysis (50% to 99%).[12] Although 10 times as many patients treated with thrombolysis had significant or complete clearing compared with anticoagulation, less than half of the lytic patients had a good-to-excellent phlebographic outcome.[12] Furthermore, systemic thrombolysis for DVT is associated with a greater risk of bleeding compared with that observed with anticoagulation alone.[40] These results led to the introduction of CDT to minimize bleeding and enhance clot dissolution.

Catheter-Directed Thrombolysis

Numerous studies have demonstrated good outcomes of CDT for acute DVT.[11] Three large reports ($n = 422$) documented a success rate of about 80%.[28,29,41] If

Numerous studies have demonstrated good outcomes of catheter-directed thrombolysis for acute DVT. Three large reports ($n = 422$) documented a success rate of about 80%.

treatment had been restricted only to patients with acute iliofemoral DVT, the initial success rate of 80% likely would have been higher, because clinicians treated patients with more chronic and extensive thrombus, resulting in somewhat lower overall success rates.[12] The largest published experience with CDT has come from the NVTR. The NVTR studies demonstrated successful removal of thrombus and improved quality of life in patients successfully treated by CDT.[12,28,41]

Open Surgical Thrombectomy

Open iliofemoral venous thrombectomy with ligation of the femoral vein was the treatment of choice for DVT in the 1960s.[42,43] Femoral vein ligation was performed to prevent subsequent pulmonary embolism and thrombectomy was performed to improve venous outflow from the leg.[6] Currently, surgical thrombectomy is typically reserved for patients with contraindications to pharmacological thrombolysis or for those in whom other modalities have failed. Inferior vena cava filters now serve the same purpose that femoral vein ligation did in the past.

Percutaneous Mechanical Thrombectomy

Percutaneous mechanical thrombectomy was the next technology directed at treatment of acute DVT because it presents the potential advantage of rapid clearance of thrombus from occluded venous segments.[44] The thrombus is briskly cleared either with a rotational or hydrodynamic mechanism.[44,45]

Rotational thrombectomy devices such as the Amplatz thrombectomy device (Microvena), the Arrow-Trerotola percutaneous thrombolytic device (Arrow), and the Cragg-Castaneda thrombolytic brush (Micro Therapeutics) use a high-speed rotating basket or impeller to macerate the thrombus.[6] The resultant small particles travel to the pulmonary arterial circulation.[45–47] Rotational, or "wall-contacting," devices can potentially damage the endothelial lining of the vein.[6] Although no clinical data is yet available, the Bacchus Fino device (Bacchus Vascular) is designed using a rotating Archimedes screw that theoretically may prevent wall contact by a helically oriented nitinol framework.[6] The fragmented thrombus is then extracted into a sheath through rotational motion.[6]

Hydrodynamic, or "rheolytic," recirculation devices such as the AngioJet (Possis Medical), the Hydrolyzer (Cordis), and the Oasis Thrombectomy System (Boston Scientific) are based on the Venturi effect.[6] The high-speed jets are directed retrograde into the thrombus and the material is then aspirated into the device.[6] A study of 37 patients reported >50% thrombus extraction in 59% of patients treated with these devices and symptomatic improvement in 82%.[45,46]

Pharmacomechanical Thrombectomy

Except in patients with bleeding diatheses, mechanical thrombectomy devices may be used in combination with pharmacological thrombolysis.[45,46] Using these 2 treatment modalities simultaneously may offer the best opportunity for rapid clearance of thrombus. This could decrease the duration and dose of the thrombolytic agent.[6] The Bacchus Trellis device (Bacchus Vascular) uses "isolated thrombolysis" to deliver and disperse small bolus doses of a thrombolytic drug into the venous thrombus. High concentrations of thrombolytic agents are injected through a catheter with proximal and distal occlusion balloons. The occlusive balloons are designed to limit leakage of thrombolytic agent into the systemic circulation, potentially reducing the risk of bleeding complications, while a sheath is designed to aspirate contents between the balloons.[6] The central balloon is intended to reduce embolization of particulate debris to the pulmonary circulation.[6] A sinusoidal nitinol wire placed within the catheter is rotated to mix the blood between the balloons.[6] The Trellis device has been used with some success to rapidly remove thrombus in patients with DVT.[47] Although promising, experience with this type of apparatus is limited and there are few published reports documenting its usefulness in patients with acute DVT. A recent registry presentation of 771 patients demonstrated 97% technical success at achieving >50% thrombus removal in acute DVT, with 54% achieving >95%

> *Except in patients with bleeding diatheses, mechanical thrombectomy devices may be used in combination with pharmacological thrombolysis. Using these 2 treatment modalities simultaneously may offer the best opportunity for rapid clearance of thrombus, as well as decrease the duration and dose of the thrombolytic agent.*

clot removal.[48,49] In addition, 84% of acute DVT cases were treated in a single session (Fig. 2).[48,49]

"Powerpulse" refers to the use of the AngioJet Rheolytic Thrombectomy System (Possis) to provide single-session pharmacomechanical CDT of DVT by forcefully pulse-spraying a thrombolytic drug into the venous thrombus to thereby achieve better drug dispersion and faster treatment. A similar registry has been presented regarding this technique where 160 patients were studied. Eighty-five percent of cases were completed in less than 1 day.[48,49] Seventy-three percent of venous segments treated were widely patent with no residual thrombus at the completion of the procedure (Fig. 3).[48,49]

These 2 techniques have revolutionized the treatment of acute DVT with the prospect of allowing endovascular treatment of DVT to be performed as an outpatient with little or no morbidity.

> *Two new techniques– isolated thrombolysis involving a trellis device and single-session thrombolysis involving a powerpulse jet– have revolutionized the treatment of acute DVT with the prospect of allowing endovascular treatment of DVT to be performed as an outpatient with little or no morbidity.*

Venous Angioplasty and Stenting

Venous angioplasty and/or stenting after CDT has been performed in 3 situations: flow-limiting iliocaval stenoses; flow-limiting femoropopliteal stenoses; and short-segment, residual infrarenal iliocaval thrombosis.[2] Stents are often preferred over balloon angioplasty, particularly in the iliac veins, because the fibrotic nature of iliac venous stenosis can sometimes result in elastic coil after balloon angioplasty.[50] Angioplasty is more often used in vessels that cross joints, such as the popliteal vein.[29]

A specific disease entity that can be treated with percutaneous venoplasty and stenting is May-Thurner syndrome, which typically affects female patients in the second to fourth decade.[51,52] May-Thurner syndrome is due to thickening and intimal hyperplasia of the left common iliac vein caused by the arterial pressure of the crossing right common iliac artery.[53] This ultimately results in stenosis and venous thrombosis.

FUTURE PERSPECTIVES

The Acute Venous Thrombosis: Thrombus Removal with Adjunctive Catheter-Directed Thrombolysis (ATTRACT) trial is a phase III, randomized, controlled trial involving 47 centers and funded by the National Heart, Lung and Blood Institute. It is analyzing 692 patients with acute, symptomatic proximal DVT.[54] Pharmacological CDT with standard therapy versus standard therapy alone will be compared. Exclusion criteria include patients with a high bleeding risk, intracranial lesions, phlegmasia, massive pulmonary emboli, symptom duration >2 weeks, same-leg DVT within the last 2 years, active cancer, and patients who are non-ambulatory or who have a short life expectancy.[54] Enrollment has begun at 28 centers in the United States. Investigators are from a variety of subspecialties, including radiology, surgery, internal medicine, emergency medicine, statistics, and economics. Dr. Suresh Vedantham of the Mallinckrodt Institute of Radiology at Washington University School of Medicine in St. Louis, MO, is the principal investigator.

The primary endpoint of the ATTRACT Trial is the occurrence of PTS at 24 months follow-up (by the Villalta PTS scale).[54] A secondary endpoint is the evaluation of PTS severity at 6, 12, 18, and 24 months using the Villalta scale, the clinical-etiology-anatomy-pathophysiology (CEAP) severity score, and the venous clinical severity score (VCSS). The ATTRACT Trial will also evaluate leg pain (Likert scale), swelling (calf circumference), and early quality of life (short-form [SF] health survey 36, VEINES quality of life/symptoms [QOL/Sym] measure) at 10 and 30 days.[54] Additional general (SF-36) and disease-specific (VEINES-QOL/Sym) measures will be done at 6, 12, 18, and 24 months.[54] Side effects and complications, such as major bleeding, intracranial hemorrhage, pulmonary emboli, recurrent venous thromboemboli, and death, will be assessed at 10 days and 2 years.[54] Economic outcomes will also be evaluated. The ATTRACT Trial will quantitatively assess thrombus burden before and after pharmacological CDT and allow correlations to be made. Finally, an ATTRACT ultrasound substudy (n = 142) will compare reflux rates and determine if reflux predicts PTS.[54]

The American College of Chest Physicians published its first consensus statements on antithrombotic therapy in 1986 and most recently in 2008. The guidelines recommend that CDT should be offered as a treatment option to selected patients (symptoms ≤ 14 days, low bleeding risk, good functional status, and >1 year life expectancy) with extensive acute proximal DVT, particularly when the iliac vessel is involved or when limb loss is possible because of circulatory compromise.[3,55] After successful CDT, correction of underlying venous lesions using balloon angioplasty and stents should be considered.[55] Pharmacomechanical thrombolysis may be preferred to CDT because pharmacomechanical thrombolysis is associated with shorter treatment times, shorter intensive care unit and hospital stays, and reduced costs.[55] Systemic thrombolytic therapy may be used to reduce acute symptoms and post-thrombotic morbidity if CDT is not available.[55] Anticoagulation therapy with the same intensity and duration should be given after CDT to prevent the risk of recurrent venous thromboemboli.[55]

CONCLUSION

The impact of DVT, pulmonary embolus, and PTS on public health is profound, and a multidisciplinary approach is necessary to address these challenges. Anticoagulation, compression stockings, surgical thrombectomy, angioplasty/stenting, pharmacomechanical thrombectomy, and CDT are important treatment options that all have a role in the treatment of DVT. The potential of clot-removal treatments to improve long-term outcomes by preventing PTS is well documented and further research is ongoing. Physicians should improve their DVT education, collaborate with other physicians to determine the best treatment options, and support the various ongoing clinical trials that focus on the benefits of clot removal.

DISCLOSURES

Potential conflict of interest: Dr. Lookstein is a consultant for Medrad Interventional/ Possis and is on the speakers bureau for Cordis Cardiac and Vascular Institute. Dr. Giordano has no conflicts.

REFERENCES

1. Vedantham S, Grassi CJ, Ferral H, et al. For Technology Assessment Committee of the Society of Interventional Radiology. Reporting standards for endovascular treatment of lower extremity deep vein thrombosis. *J Vasc Interv Radiol* 2006; 17: 417–434.
2. Vedantham S, Thorpe PE, Cardella JF, et al. Quality improvement guidelines for the treatment of lower extremity deep vein thrombosis with use of endovascular thrombus removal. *J Vasc Interv Radiol* 2009; 20(7 suppl): S227–S239.
3. Houman Fekrazad M, Lopes RD, et al. Treatment of venous thromboembolism: guidelines translated for the clinician. *J Thromb Thrombolysis* 2009; 28: 270–275.
4. Kearon C. Natural history of venous thromboembolism. *Circulation* 2003; 107(23 suppl 1): I22–I30.
5. Haines ST. Venous thromboembolism: pathophysiology and clinical presentation. *Am J Health Syst Pharm* 2003; 60(22 suppl 7): S3–S5.
6. Augustinos P, Ouriel K. Invasive approaches to treatment of venous thromboembolism. *Circulation* 2004; 110(9 suppl 1): I27–I34.
7. Breddin HK, Hach-Wunderle V, Nakov R, et al. Effects of a low-molecular-weight heparin on thrombus regression and recurrent thromboembolism in patients with deep-vein thrombosis. *N Engl J Med* 2001; 344: 626–631.
8. Labas P, Ohradka B, Vladimir J, et al. The home treatment of deep vein thrombosis with low molecular weight heparin, forced mobilisation and compression. *Int Angiol* 2000; 19: 303–307.
9. Kahn SR, Ginsberg JS. Relationship between deep venous thrombosis and the post-thrombotic syndrome. *Arch Intern Med* 2004; 164: 17–26.
10. Kahn SR, Ginsberg JS. The post-thrombotic syndrome: current knowledge, controversies, and directions for future research. *Blood Rev* 2002; 16: 155–165.
11. Shull KC, Nicolaides AN, Fernandes J, et al. Significance of popliteal reflux in relation to ambulatory venous pressure and ulceration. *Arch Surg* 1979; 14: 1304–1306.
12. Comerota AJ, Gravett MH. Iliofemoral venous thrombosis. *J Vasc Surg* 2007; 46: 1065–1076.
13. Johnson BF, Manzo RA, Bergelin RO, Strandness DE Jr. Relationship between changes in the deep venous system and the development of the post-thrombotic syndrome after an acute episode of lower limb deep vein thrombosis: a one- to six-year follow-up. *J Vasc Surg* 1995; 21: 307–313.
14. Strandness DE Jr., Langlois Y, Cramer M, et al. Long-term sequelae of acute venous thrombosis. *JAMA* 1983; 250: 1289–1292.
15. Beyth RJ, Cohen AM, Landefeld CS. Long-term outcomes of deep-vein thrombosis. *Arch Intern Med* 1995; 155: 1031–1037.
16. Meissner MN, Strandness DE. Pathophysiology and natural history of deep venous thrombosis. In: Rutherford RB, ed. *Vascular Surgery*. 5th ed. Philadelphia, PA: WB Saunders; 2000; 1920–1937.
17. Douketis JD, Crowther MA, Foster GA, Ginsberg JS. Does the location of thrombosis determine the risk of disease recurrence in patients with proximal deep vein thrombosis? *Am J Med* 2001; 110: 515–519.
18. Delis KT, Bountouroglou D, Mansfield AO. Venous claudication in iliofemoral thrombosis: long-term effects on venous hemodynamics, clinical status, and quality of life. *Ann Surg* 2004; 239: 118–126.
19. Bergqvist D, Jendteg S, Johansen L, et al. Cost of long-term complications of deep venous thrombosis of the lower extremities: an analysis of a defined patient population in Sweden. *Ann Intern Med* 1997; 126: 454–457.
20. Heit JA, Rooke TW, Silverstein MD, et al. Trends in the incidence of venous stasis syndrome and venous ulcer: a 25-year population-based study. *J Vasc Surg* 2001; 33: 1022–1027.

21. Cho JS, Martelli E, Mozes G, et al. Effects of thrombolysis and venous thrombectomy on valvular competence, thrombogenicity, venous wall morphology, and function. *J Vasc Surg* 1998; 28: 787–799.
22. Rhodes JM, Cho JS, Gloviczki P, et al. Thrombolysis for experimental deep venous thrombosis maintains valvular competence and vasoreactivity. *J Vasc Surg* 2001; 31: 1193–1205.
23. Killewich LA, Bedford GR, Beach KW, Strandness DE Jr. Spontaneous lysis of deep venous thrombi: rate and outcome. *J Vasc Surg* 1989; 9: 89–97.
24. Markel A, Manzo RA, Bergelin RO, Strandness DE Jr. Valvular reflux after deep vein thrombosis: incidence and time of occurrence. *J Vasc Surg* 1992; 15: 377–382.
25. Meissner MH, Manzo RA, Bergelin RO, et al. Deep venous insufficiency: the relationship between lysis and subsequent reflux. *J Vasc Surg* 1993; 18: 596–605.
26. Plate G, Eklof B, Norgren L, et al. Venous thrombectomy for iliofemoral vein thrombosis–10-year results of a prospective randomised study. *Eur J Vasc Endovasc Surg* 1997; 14: 367–374.
27. Plate G, Akesson H, Einarsson E, et al. Long-term results of venous thrombectomy combined with a temporary arterio-venous fistula. *Eur J Vasc Surg* 1990; 4: 483–489.
28. Bjarnason H, Kruse JR, Asinger DA, et al. Iliofemoral deep venous thrombosis: safety and efficacy outcome during 5 years of catheter-directed thrombolytic therapy. *J Vasc Interv Radiol* 1997; 8: 405–418.
29. Mewissen MW, Seabrook GR, Meissner MH, et al. Catheter-directed thrombolysis for lower extremity deep venous thrombosis: report of a national multicenter registry. *Radiology* 1999; 211: 39–49.
30. Comerota AJ, Paolini D. Treatment of acute iliofemoral deep venous thrombosis: a strategy of thrombus removal. *J Vasc Surg* 2007; 45: 640.
31. Porter JM, Rutherford RB, Clagett GP, et al. Reporting standards in venous disease. *J Vasc Surg* 1988; 8: 172–181.
32. Strandness DE, Langlois Y, Cramer M, et al. Long-term sequelae of acute venous thrombosis. *JAMA* 1983; 250: 1289–1292.
33. Semba CP, Dake MD. Iliofemoral deep venous thrombosis: aggressive therapy with catheter-directed thrombolysis. *Radiology* 1994; 191: 487–494.
34. Markel A, Manzo RA, Bergelin RO, Strandness DE, Jr. Valvular reflux after deep vein thrombosis: incidence and time of occurrence. *J Vasc Surg* 1992; 15: 377–384.
35. Meissner MH, Manzo RA, Bergelin RO, et al. Deep venous insufficiency: the relationship of lysis and subsequent reflux. *J Vasc Surg* 1993; 18: 596–608.
36. Brown WD, Goldhaber SZ. How to select patients with deep vein thrombosis for tPA therapy. *Chest* 1989; 95(5 suppl): 276S–278S.
37. Blinc A, Kennedy SD, Bryant RG, et al. Flow through clots determines the rate and pattern of fibrinolysis. *Thromb Haemost* 1994; 71: 230–235.
38. Blinc A, Francis CW. Transport processes in fibrinolysis and fibrinolytic therapy. *Thromb Haemost* 1996; 76: 481–491.
39. Goldhaber SZ, Buring JE, Lipnick RJ, et al. Pooled analyses of randomized trials of streptokinase and heparin in phlebographically documented acute deep venous thrombosis. *Am J Med* 1984; 76: 393–397.
40. Comerota AJ, Kagan SA. Catheter-directed thrombolysis for the treatment of acute iliofemoral deep venous thrombosis. *Phlebology* 2001; 15: 149–155.
41. Keisker HW, Bowers RF. Results obtained by superficial femoral vein ligation. *Surgery* 1960; 47: 224–229.
42. Agrifoglio G, Edwards EA. Venous stasis after ligation of femoral veins or inferior vena cava. *JAMA* 1961; 178: 1–7.
43. Frisoli JK, Sze D. Mechanical thrombectomy for the treatment of lower extremity deep vein thrombosis. *Tech Vasc Interv Radiol* 2003; 6: 49–52.
44. Sharafuddin MJ, Sun S, Hoballah JJ, et al. Endovascular management of venous thrombotic and occlusive diseases of the lower extremities. *J Vasc Interv Radiol* 2003; 14: 405–423.
45. Tadavarthy SM, Murray PD, Inampudi S, et al. Mechanical thrombectomy with the Amplatz device: human experience. *J Vasc Interv Radiol* 1994; 5: 715–724.
46. Kasirajan K, Gray B, Ouriel K. Percutaneous AngioJet thrombectomy in the management of extensive deep venous thrombosis. *J Vasc Interv Radiol* 2001; 12: 179–185.
47. Ramaiah V, Del Santo PB, Rodriguez-Lopez JA, et al. Trellis thrombectomy system for the treatment of iliofemoral deep venous thrombosis. *J Endovasc Ther* 2003; 10: 585–589.
48. Lookstein RA, Blitz L, Simoni E. Midterm results of the PEARL registry for the management of deep vein thrombosis. Paper presented at: Radiological Society of North America Annual Scientific Meeting; November 29-December 4, 2009; Chicago, IL.
49. O'Sullivan GJ, Rahzavi MK. An endovascular approach to deep venous thrombosis utilizing isolated thrombolysis and adjunctive measures. Paper presented at: Society of Interventional Radiology 33rd Annual Scientific Meeting; March 15–20, 2008; Washington, DC.
50. Semba CP, Dake MD. Catheter-directed thrombolysis for iliofemoral venous thrombosis. *Semin Vasc Surg* 1996; 9: 26–33.
51. Patel NH, Stookey KR, Ketcham DB, et al. Endovascular management of acute extensive iliofemoral deep venous thrombosis caused by May-Thurner syndrome. *J Vasc Interv Radiol* 2000; 11: 1297–1302.
52. Lamont JP, Pearl GJ, Patetsios P, et al. Prospective evaluation of endoluminal venous stents in the treatment of the May-Thurner syndrome. *Ann Vasc Surg* 2002; 16: 61–64.
53. Ferris EJ, Lim WN, Smith PL, et al. May-Thurner syndrome. *Radiology* 1983; 147: 29–31.
54. Vedantham S. Acute Venous Thrombosis: Thrombus Removal with Adjunctive Catheter-Directed Thrombolysis (ATTRACT) trial. PowerPoint presentation. http://www.sirfoundation.org/clinical-trials/attract.shtml. Accessed October 2009.
55. Kearon C, Kahn SR, Agnelli G, et al. Antithrombotic therapy for venous thromboembolic disease: American College of Chest Physicians evidence-based clinical practice guidelines (8th ed) [published correction appears in *Chest* 2008;134:892]. *Chest* 2008; 133(6 suppl): 454S–545S.

Mount Sinai Journal of Medicine 77:296–303, 2010

Visceral Artery Aneurysms: Review of Current Management Options

Ulka Sachdev-Ost, MD

University of Pittsburgh Medical Center, Pittsburgh, PA

OUTLINE

Splenic Artery Aneurysms
Hepatic Artery Aneurysms
Superior Mesenteric Artery Aneurysms
Celiac Artery Aneurysms
Gastroduodenal, Pancreaticoduodenal, and Inferior Mesenteric Artery Aneurysms
Visceral Artery Pseudoaneurysms
Conclusion

ABSTRACT

Visceral artery aneurysms are relatively rare clinical entities, although their detection is rising due to an increased use of cross-sectional imaging. Rupture is the most devastating complication, and is associated with a high morbidity and mortality. For this reason, elective repair is preferable in the appropriately chosen patient. In general, splenic artery aneurysms measuring 2 cm or larger and those found in women of childbearing age and in persons undergoing liver transplantation should be treated. Hepatic artery aneurysms 2 cm or larger and those that are multiple or nonatherosclerotic in nature should be repaired in the appropriate patient due to a higher risk of rupture. Endovascular coil embolization has excellent success rates and is the first-line treatment for anatomically suitable splenic artery aneurysms and intrahepatic hepatic artery aneurysms. However, reperfusion is an important complication of endovascular management. Aneurysms involving the celiac, superior mesenteric, pancreaticoduodenal, gastroduodenal, and inferior mesenteric arteries, as well as visceral artery pseudoaneurysms, are unpredictable and should be repaired in the appropriate medical patient. These aneurysms are often amenable to ligation due to the presence of collateral circulation. Endovascular management is particularly useful in the treatment of pseudoaneurysms where comorbidities and previous surgery make open surgical repair less desirable. ***Mt Sinai J Med 77:296–303, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: aneurysm, endovascular coil embolization, splanchnic circulation, visceral artery.

Address Correspondence to:

Ulka Sachdev-Ost
Department of Surgery
University of Pittsburgh Medical Center
Pittsburgh, PA
Email: sachdevu2@upmc.edu

Visceral artery aneurysms (VAA) are uncommon entities that affect the splenic, hepatic, celiac, and superior mesenteric arteries. Increased imaging of the abdomen with ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI) to evaluate nonvascular pathology has led to a concomitant increase in their diagnosis in asymptomatic patients. Elective repair is recommended in certain situations to prevent the high morbidity and mortality that is associated with aneurysm rupture.

Conventional repair of VAA has included excision, bypass, ligation, and end-organ resection. Endovascular techniques have gained popularity and have become the first-line treatment for certain types of aneurysms, such as those that affect the mid- to proximal splenic artery. Finally, laparoscopic techniques have been used with excellent results, and have allowed for the treatment of patients who would otherwise be unfit for open surgery due to comorbidities. This article will systematically review the natural history (when known), management

> *Endovascular techniques have gained popularity and have become the first-line treatment for certain types of aneurysms, such as those that affect the mid- to proximal splenic artery.*

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20181

© 2010 Mount Sinai School of Medicine

recommendations, and current treatment options for aneurysms involving branches of the celiac and superior mesenteric arteries, with particular emphasis on endovascular therapies. Management of visceral artery pseudoaneurysms also will be discussed.

SPLENIC ARTERY ANEURYSMS

Splenic artery aneurysms (SAA) account for about 50% to 75% of all VAA,[1–6] are estimated to have a prevalence of <0.1%, and may be associated with other intra-abdominal aneurysms involving visceral (3%) and renal (14%) arteries.[4] They have a female to male predominance of 3–4:1[3,4,7] and are associated with multiparity.[3] In a review of 217 patients with SAA, mean age was 61 years, 79% were women, 95% of the aneurysms were solitary, and the majority were asymptomatic.[3] However, rupture was the presenting symptom in close to 5% of patients, with a mean diameter of 3.5 cm (3.2 cm for males, 2.3 cm for females).[3] In contrast, the mean diameter of the nonruptured SAAs was 2.2 cm. Interestingly, calcification, which is often thought to offer evidence of chronicity and stability, was seen in 90% of the ruptured aneurysms,[3] and should therefore not be used to guide management decisions.

True SAAs are associated with multiparity as well as portal hypertension.[4,5,7] Hormonal changes and increased portal congestion associated with pregnancy are thought to lead to loss of structural integrity within the splenic artery, leading to aneurysm formation.[4] SAA has a prevalence as high as 10% in patients with portal hypertension,[4] and aneurysm rupture after liver transplantation is not an uncommon outcome. In a series of 30 patients who underwent repair of SAA at a busy liver transplant center, 20 patients had a previous orthotopic liver transplant (OLT) and 12 patients had a history of portal hypertension. Rupture was an indication for repair in almost half the patients in this study, and occurred within 2 weeks of liver transplantation in 7 patients.[7] Other etiologies of SAA include systemic hypertension, as well as less-common conditions such as medial fibrodysplasia and $\alpha - 1$ antitrypsin deficiency.[4] Although calcification is commonly seen within them, atherosclerosis does not appear to be a major factor in the development of SAA.[4]

The natural history of untreated SAA is poorly defined, but may be best described in the Mayo Clinic experience by Abbas *et al.* in which the decision was made to observe 168 patients at the initial presentation of SAA.[3] Sixty-five percent of these patients had aneurysm size documented on presentation, and 47% had size information available at follow-up. Mean aneurysm size at presentation in the nonoperative group was 2.1 cm, with approximately one-third of those patients having aneurysms between 2 and 3 cm. Of the 79 patients in whom follow-up measurements were obtained, 11 (14%) had evidence of aneurysm growth; of these, 3 underwent elective repair. None of the patients who were observed ruptured their aneurysms during a mean follow-up period of 75 months (range, 1–371 months).[3] To summarize natural history information as well as indications of repair, SAA can have an appreciable growth rate, and increased aneurysm size is associated with rupture. Current available data would indicate that repair should be offered for symptomatic aneurysms, aneurysms ≥2 cm aneurysms found in women of childbearing age (because rupture in pregnant women is associated with high maternal and fetal mortality),[8] and in patients with SAA who are undergoing OLT. In the latter case, repair of the aneurysm at the time of transplant has been reported.[7]

Treatment options for SAA include open surgical repair, endovascular management with either coils or covered stents, and, more recently, laparoscopic excision. The location of the aneurysm as well as its presentation dictates the type of open procedure that is performed. Resection with end-end repair can be performed in many cases, owing to the predominantly proximal location of SAA and to the redundancy and tortuosity of the artery.[5] This allows for splenic preservation, which has important implications for the immune system. However, splenectomy may become necessary in aneurysms involving the splenic hilum,[9] and is performed more commonly in the setting of aneurysm rupture.[3,5,7]

Endovascular management of SAA has gained popularity over the last decade, owing to its acceptable technical success rate and low morbidity.[1,2,6,10,11] Percutaneous embolization with either coils or glue (N-butyl cyanoacrylate) is particularly appropriate for saccular aneurysms, fusiform aneurysms with good collateral flow to the end organ, and aneurysms of vessels in which other sources of arterial flow to the end organ are present.[2,10] Thus, SAA are often ideal candidates for coil embolization (Figure 1). Triaxial systems that utilize microcatheters allow for super-selective catheterization of branches to optimize aneurysm exclusion.[2,10] Procedures can be performed under local anesthesia with sedation and are generally ambulatory in the elective setting.[1] Percutaneous access is most commonly achieved through cannulation of the femoral artery.[2,10] When there is no evidence of rupture, systemic heparin is administered to prevent thrombosis of the sheath and access vessels. Rates of successful exclusion of SAA range

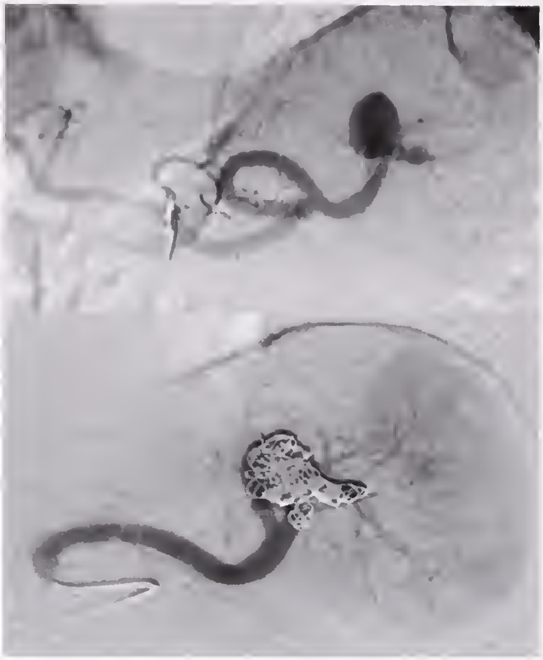


Fig 1. Angiogram demonstrating splenic artery aneurysm (A) before and (B) after successful coil embolization.

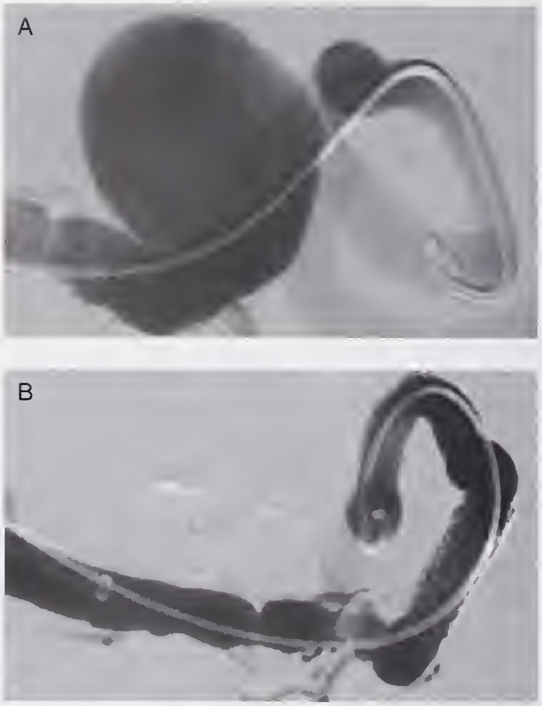


Fig 2. Angiogram showing (A) splenic artery aneurysm and (B) successful stent graft exclusion.

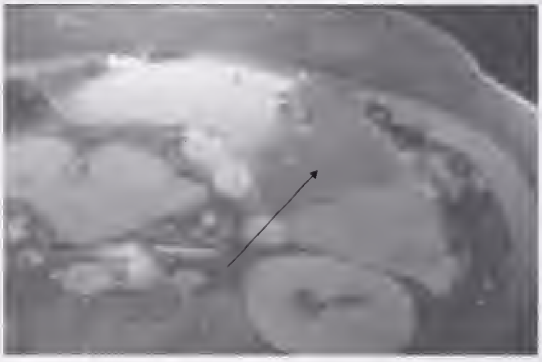


Fig 3. Arrow demonstrates splenic infarct after successful splenic artery embolization. This patient had splenomegaly associated with portal hypertension.

from 90% to 100%.[1,2,11,12] Placement of stent grafts to maintain perfusion in the main artery has also been described with success in the splenic artery[1,13] (Figure 2); however, tortuosity of the artery can complicate stent graft placement and deployment.

Complications of SAA coil embolization include splenic infarct (Figure 3) and reperfusion of the aneurysm, which can occur in 5% to 20% of patients.[1,12] Thus, yearly follow-up with CT or MRI is necessary to evaluate for leak and subsequent growth.[12] In our previously reported experience at Mount Sinai Medical Center, 3 out of 15 patients treated with endovascular coil embolization developed evidence of reperfusion. However, all were able to be treated by repeat embolization through endoluminal access, avoiding the need for laparotomy.[1] Splenic infarcts have been seen more commonly in patients with portal hypertension and hilar location of aneurysms, but usually are successfully treated with pain control.[1,2,6,14] In reported series, splenic infarcts occurred in 25% to 40% of the patients treated, yet no patient developed sequlae requiring further management.[1,2,12,14]

Laparoscopic management of SAA was first reported in 1993,[15] and though it is not a commonly performed procedure, it does offer an alternative to endovascular and open surgical management. Although the literature describes only small series, outcomes appear to be acceptable, and the procedure offers certain advantages over other treatment options. Patients are spared the need for yearly CT follow-up to evaluate for aneurysm reperfusion, which is required for endovascular management and may be problematic in young patients. Furthermore, splenectomy can be performed at the time of the procedure, which obviously prevents the syndrome associated with postembolization infarction, pain, and potential for infection. In 1 small series, splenectomy was performed when laparoscopic-assisted Doppler sonography demonstrated markedly diminished flow to the spleen after aneurysm exclusion.[16] This occurred most commonly in hilar aneurysms.

Although, theoretically, end-to-end arterial anastomoses can be performed laparoscopically to maintain arterial perfusion, in the aforementioned small series a robotically assisted anastomosis was performed in the 1 case in which reconstruction was performed, adding to the complexity and cost of the procedure.[16]

HEPATIC ARTERY ANEURYSMS

Hepatic artery aneurysms (HAA) are the second most common true VAA (Figure 4). However, increased performance of percutaneous biliary procedures, nonoperative management of trauma, and liver transplantations has led to an increase in hepatic artery pseudoaneurysms,[1,2,9,10,17,18] making them the most commonly reported VAA in the period from 1985 to 1995.[17,19] In contrast to SAA, HAA are more commonly identified in men (male/female [M/F] ratio 3:2),[4,20] although the mean age of presentation is similar (sixth to seventh decade).[3,4,20] In a series of 36 true HAA reported by the Mayo Clinic, hypertension was the most commonly associated comorbidity, occurring in 72% of patients.[20] Nonhepatic VAA were identified with HAA in 31% of patients and were identified in the splenic artery in 20% of patients.[20] The majority of aneurysms were solitary and atherosclerotic in nature, and aneurysm rupture was seen only in patients with multiple aneurysms or with aneurysms of nonatherosclerotic etiology (eg, fibromuscular dysplasia, polyarteritis nodosa, or mycotic).[20] Most aneurysms were located in the extrahepatic circulation, and fewer were seen in both the extrahepatic and intrahepatic arterial system. Pure intrahepatic aneurysms were an uncommon finding.[20]

> *Increased performance of percutaneous biliary procedures, nonoperative management of trauma, and liver transplantations has led to an increase in hepatic artery pseudoaneurysms, making them the most commonly reported visceral artery aneurysm in the period from 1985 to 1995.*

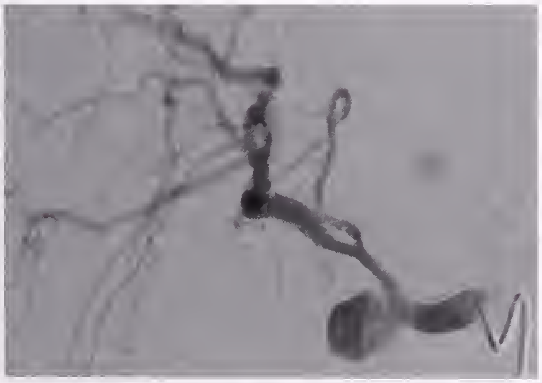


Fig 4. Angiogram demonstrating hepatic artery aneurysm.

Most HAA are identified as incidental findings on cross-sectional imaging or abdominal ultrasound.[4] However, abdominal or back pain can be present in rapidly expanding aneurysms.[4] Importantly, rupture into the biliary tree is more common than free, intraperitoneal rupture, and can present with jaundice, biliary colic, and gastrointestinal hemorrhage as is described in Quincke's triad.[4,19] In the Mayo Clinic experience, 22 patients (61%) were managed nonoperatively for aneurysms <2.0 cm (n = 14) and >3.0 cm (n = 4).[20] Two of these patients presented ruptured, and 1 died. The other had contained hematoma, was observed, and was alive at 3 months follow-up. In 3 patients managed nonoperatively, aneurysm growth was noted, but repair was precluded due to either poor medical condition or aneurysm size <2.5 cm.[20] None of the observed patients, including those with aneurysm diameter >3 cm, proceeded to rupture.[20]

Based on the Mayo Clinic experience, which represents one of the largest series of true HAA in the literature, indications for treatment include symptomatic aneurysms, nonatherosclerotic aneurysms, multiple aneurysms, and aneurysms >2 cm in good-risk patients with a life expectancy of at least 2 years. In their series, patients with significant comorbidities who presented with aneurysms >3 cm were observed, and none proceeded to rupture.[20] Alternatives to open surgical management, including endovascular treatment, may allow for treatment of large aneurysms in the medically compromised patient.[18]

Traditional surgical management of HAA depends on the location of the aneurysm, presence of collateral flow, and medical state of the patient. Procedures can include ligation, ligation with bypass using autologous vein or prosthetic graft, hepatectomy, and, in rare instances, OLT.[1,19,20] In either the open or endovascular strategy, ligation or coil embolization of the hepatic artery should be performed only if the portal vein is patent.[18] Aneurysms involving the proper hepatic artery require vascular reconstruction as a component of repair, whereas common HAA may be managed without reconstruction if collateral flow through the gastroduodenal artery is adequate.[4,5,17,18] Complications of open surgical repair include graft thrombosis, bile leak, and intra-abdominal abscess or sepsis.[1,20] Open surgical repair for rupture is associated with a high morbidity and mortality (100% and 33%, respectively[20]); thus, elective repair, when appropriate, is preferred.

Literature describing the use of endovascular techniques to manage HAA is limited to small series. Percutaneous embolization is considered first-line treatment for intrahepatic aneurysms owing to the complicated nature of open repair.[21] Percutaneous approaches appear to have particular use in the treatment of hepatic artery pseudoaneurysms, in which previous abdominal surgery and medical comorbidity are prominent features.[1,10,17] In our series describing the experience at Mount Sinai, an endovascular approach predominated in the treatment of hepatic artery pseudoaneurysms, whereas true aneurysms involving the extrahepatic circulation were more likely to be treated by surgical resection (n = 10:1, pseudo:true treated endoluminally versus 4:7, pseudo:true treated by open surgery; $P = 0.002$).[1] In another series from the Cleveland Clinic, technical success at coil-embolizing 10 HAA was 80%, and failures were attributed to an inability to cannulate the aneurysm neck.[10] Although follow-up imaging was not available for all patients, no reperfusion was seen in patients in whom imaging was available.[10] In another small series of 12 hepatic artery embolizations, no evidence of hepatic ischemia occurred after aneurysm ablation with either coils or glue.[2]

> *Percutaneous embolization is considered first-line treatment for intrahepatic aneurysms owing to the complicated nature of open repair.*

Stent graft repair may be desirable in patients with baseline hepatic insufficiency, or in aneurysms involving the proper hepatic artery in order to maintain antegrade perfusion. However, anatomic suitability for stent graft repair must be determined preoperatively to evaluate the adequacy of the seal zones, the potential for kinking of the artery, and the accessibility of proximal vessels for delivery sheaths. Reported complications of endovascular therapy include access-related problems such as groin hematomas, pseudoaneurysms, and arterial thrombosis, which can occur during any percutaneous intervention. Procedure-specific complications can also include arterial dissection[21] and aneurysm rupture.

SUPERIOR MESENTERIC ARTERY ANEURYSMS

Superior mesenteric artery (SMA) aneurysms are the third most common type of VAA, and are a rare entity. Most commonly they involve the proximal 5 cm of the SMA.[22] Extension of aneurysm-related thrombus or dissection beyond the collateral circulation of the celiac and inferior mesenteric arteries may result in ischemic symptoms.[23] Historically, the etiology of SMA aneurysms was infectious.[22] However, in a series of 21 patients, etiologies varied from atherosclerosis and cystic medial dysplasia to collagen vascular disorders and polyarteritis nodosa.[24] Medial dysplasia may be a secondary finding.[23] Only 1 patient had evidence of infection. In that same series, 8 patients presented ruptured, and nearly half of the patients with rupture subsequently died.[24] In contrast to SAA aneurysms, patients with calcified SMA aneurysms did not have rupture either at presentation or follow-up.[24] There was no accurate description of aneurysm size in patients who ruptured, but the mean size in patients who did not rupture was 2.2 cm. Furthermore, male patients were more prone to rupture. Of the 13 patients who did not rupture, 10 were asymptomatic.[24] Patients were repaired unless comorbidities precluded operative intervention. Although elective repair did not require bowel resection, repair for rupture was associated with small-bowel resection in 3 out of the 8 patients for bowel ischemia.[24] Thus, elective repair may be beneficial in terms of bowel preservation. Based on limited literature, indications for repair other than rupture include size ≥2 cm in a good-risk patient, growth, and symptoms.

Ligation of the vessels that enter and exit the aneurysm is one of the more common surgical approaches to treat SMA aneurysms.[23,24] Collaterals from the celiac and inferior mesenteric arteries are usually sufficient to prevent ischemia, although test occlusion with an intraoperative assessment of bowel viability can be performed.[23] Other open surgical approaches include endoaneurysmorraphy and arterial reconstruction with bypass after ligation using either vein or prosthetic graft. Vein is the preferred conduit in the setting of an infected aneurysm. Operative exposure is transmesenteric in more distal aneurysms, whereas those in the proximal SMA may require exposure through medial visceral rotation.[23]

Reports of endovascular management of SMA aneurysms describe the use of coil embolization and stent grafts.[6,24] Percutaneous management is particularly useful in patients with a hostile abdomen and medical comorbidities. In our previously published experience, 3 patients with SMA pseudoaneurysms were successfully managed by endovascular means without procedure-related complications. Two of these patients had hostile abdomens, and 1 had medical comorbidity.[10] In general, literature on these techniques is limited to very small series or case reports, thus long-term outcomes are difficult to assess.

Reports of endovascular management of superior mesenteric artery (SMA) aneurysms describe the use of coil embolization and stent grafts. Percutaneous management is particularly useful in patients with a hostile abdomen and medical comorbidities.

CELIAC ARTERY ANEURYSMS

Celiac artery aneurysms (CAA) are quite rare, accounting for 4% of VAA.[22] The etiology of these aneurysms is usually atherosclerosis, and infectious causes are rare. Most are asymptomatic, but symptoms can mimic pancreatitis due to the location of the arteries.[23] Of note, association with other aneurysms is common, occurring in 66% of patients in 1 series.[25] Concomitant aneurysms included aortic, renal, popliteal, and femoral in that report.[25] In 1 series, CAA were the most common VAA encountered (9 of 23 patients), most likely because SAA were seen and treated by interventional radiologists.[26]

Because these aneurysms are so rare, reports of their natural history are limited. Again, the Mayo Clinic experience represents the largest compilation of patients with CAA (n = 18), some of whom were treated and some of whom were observed.[25] The patients who were observed were considered too high-risk for repair, and were thus followed with serial imaging.[25] Mean aneurysmal size was 2.1 cm (range, 1.5–3.5 cm) in this group of 8 patients who were observed. One patient with a 2.5-cm aneurysm who declined repair ruptured 5 years after presentation, whereas the others did not demonstrate any sign of growth over a 91-month mean follow-up period (range: 1–371 months).[25] Thus, observation in medically compromised patients is a reasonable option.

Traditional open repair can be performed through a transabdominal route, and may require medial visceral rotation of left colon, spleen, and pancreas to gain access to the proximal celiac artery.[23] A thoracoabdominal approach may be required in some instances when the aneurysm is large.[23] Ligation can be performed in the absence of liver pathology.[25] Aneurysmectomy with arterial reconstruction using vein or prosthetic graft is another approach, using the supraceliac aorta as an inflow source.[23] In the Mayo Clinic experience, vein bypasses occluded within 6 months follow-up. Kinking of the graft and thrombosis in the setting of competitive flow were proposed explanations for this finding.[25] Thus, prosthetic grafts may offer some advantage in this setting.

Endovascular management of CAA has been described and may be appropriate in high-risk patients without liver dysfunction and without disease of the collateral circulation, including the SMA and gastroduodenal arteries.[2,11,27] Endovascular treatment of CAA may also be performed during endoluminal repair of other VAA,[1] or before treatment of aortic aneurysms.[27]

Endovascular management of celiac artery aneurysms (CAA) has been described and may be appropriate in high-risk patients without liver dysfunction and without disease of the collateral circulation, including the SMA and gastroduodenal arteries.

The nature of the collateral circulation theoretically allows for embolization of the inflow, outflow, and branches of the aneurysm without arterial reconstruction. Though reports of this technique are limited in regard to CAA repair, celiac artery coverage and occlusion is sometimes utilized in endovascular repair of thoracic artery aneurysms (TEVAR). The safety of celiac artery coverage in that setting is a matter of current investigation. Most authors recommend preoperative angiography to evaluate collateral circulation,[28,29] and some have advocated selective arteriography through the SMA with concomitant celiac artery test occlusion to evaluate the adequacy of mesenteric collaterals.[28,30] However, in 1 study, test occlusion of the celiac and mesenteric angiography did not prevent foregut ischemia in the setting of celiac artery coverage. Furthermore, in the same study, coverage of the celiac artery during TEVAR led to foregut ischemic complications in 25% of patients who did not have preoperative celiac artery revascularization.[28] Complications included an ischemic gastric ulcer, gangrenous cholecystitis, liver abscess, and exacerbation of cirrhosis.[28] Whether the ischemic complications were due to inadequate flow or procedural embolization was an unresolved issue in that report. Complications specifically related to endovascular aneurysm exclusion do not appear to commonly involve organ ischemia,[2,11] although reports are limited. Late coil migration leading to fatal gastrointestinal hemorrhage has been described.[31]

GASTRODUODENAL, PANCREATICODUODENAL, AND INFERIOR MESENTERIC ARTERY ANEURYSMS

True aneurysms involving the gastroduodenal, pancreaticoduodenal, and inferior mesenteric arteries are rare. However, their behavior is unpredictable. In 1 series of VAA in 34 patients, all 6 patients with pancreaticoduodenal artery aneurysms presented ruptured, and 2 died.[32] Thus, aneurysms in these locations should be treated. Management with ligation or coil embolization is appropriate, as reconstruction is not only technically difficult, but is likely not required due to adequate collateralization.[11,32] However, adequate collateral flow should be documented with preoperative imaging if permitted. Imaging with CT and magnetic resonance angiography demonstrate excellent resolution for preoperative planning.[22] However, angiography may allow for better evaluation of real-time flow dynamics.

VISCERAL ARTERY PSEUDOANEURYSMS

Visceral artery pseudoaneurysms (VAPA) occur in the setting of infection, chronic inflammation, and trauma. The literature suggests that VAPA are more prone to symptoms or rupture.[2,11] In 1 series, 7 out of 25 patients who presented with rupture had VAPA secondary to chronic pancreatitis.[11] Endovascular treatment of VAPA has certain advantages in anatomically difficult locations, and in the postoperative abdomen. In our series from Mount Sinai, ruptured VAPAs were significantly more likely to be treated by endovascular techniques. Hemodynamic instability did not preclude endovascular repair, which has also been seen in other studies.[1,33] Indeed, endovascular approaches may have some advantage in the hemodynamically unstable patient with VAPA in difficult locations–angiography not only confirms the pathology, but, in addition, endovascular access may allow for rapid control with intraluminal balloon occlusion.

CONCLUSION

In summary, although VAA and VAPA are relatively rare, the mortality associated with rupture is significant. Thus, a rigorous plan should be adopted when they are recognized in the elective setting. In general, SAA measuring ≥ 2 cm and those found in women of childbearing age and in persons undergoing OLT should be treated. Hepatic artery aneurysms >2 cm and that are multiple or nonatherosclerotic in nature should be fixed in the appropriate patient due to a higher risk of rupture. Endovascular coil embolization has become the treatment of choice for anatomically suitable SAA and intrahepatic HAA. Success rates range from 85% to 100%, but reperfusion is a reported complication. Aneurysms involving the celiac and superior mesenteric arteries are much less common, but should be repaired in the appropriate medical patient. Ligation without reconstruction is feasible in the management of these aneurysms, provided that collateral circulation, often determined on preoperative imaging, is adequate. Otherwise, bypass with vein (in the setting of infection) and prosthetic graft is recommended. Finally, aneurysms involving the pancreaticoduodenal, gastroduodenal, and inferior mesenteric arteries, as well as VAPA, should be repaired, as rupture risks are high and behavior is unpredictable. Endovascular coil embolization is particularly useful in these settings, in which open operative exposure may be difficult and arterial reconstruction is not necessarily required.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Sachdev U, Baril DT, Ellozy SH, et al. Management of aneurysms involving branches of the celiac and superior mesenteric arteries: a comparison of surgical and endovascular therapy. *J Vasc Surg* 2006; 44: 718–724.

2. Tulsyan N, Kashyap VS, Greenberg RK, et al. The endovascular management of visceral artery aneurysms and pseudoaneurysms. *J Vasc Surg* 2007; 45: 276–283.
3. Abbas MA, Stone WM, Fowl RJ, et al. Splenic artery aneurysms: two decades experience at Mayo Clinic. *Ann Vasc Surg* 2002; 16: 442–449.
4. Berceli SA. Hepatic and splenic artery aneurysms. *Semin Vasc Surg* 2005; 18: 196–201.
5. Pulli R, Dorigo W, Troisi N, et al. Surgical treatment of visceral artery aneurysms: a 25-year experience. *J Vasc Surg* 2008; 48: 334–342.
6. Saltzberg SS, Maldonado TS, Lamparello PJ, et al. Is endovascular therapy the preferred treatment for all visceral artery aneurysms? *Ann Vasc Surg* 2005; 19: 507–515.
7. Lee PC, Rhee RY, Gordon RY, et al. Management of splenic artery aneurysms: the significance of portal and essential hypertension. *J Am Coll Surg* 1999; 189: 483–490.
8. Ha JF, Phillips M, Faulkner K. Splenic artery aneurysm rupture in pregnancy. *Eur J Obstet Gynecol Reprod Biol* 2009; 146: 133–137.
9. Chiesa R, Astore D, Guzzo G, et al. Visceral artery aneurysms. *Ann Vasc Surg* 2005; 19: 42–48.
10. Kasirajan K, Greenberg RK, Clair D, Ouriel K. Endovascular management of visceral artery aneurysm. *J Endovasc Ther* 2001; 8: 150–155.
11. Gabelmann A, Gorich J, Merkle EM. Endovascular treatment of visceral artery aneurysms. *J Endovasc Ther* 2002; 9: 38–47.
12. Loffroy R, Guiu B, Cercueil JP, et al. Transcatheter arterial embolization of splenic artery aneurysms and pseudoaneurysms: short- and long-term results. *Ann Vasc Surg* 2008; 22: 618–626.
13. Larson RA, Solomon J, Carpenter JP. Stent graft repair of visceral artery aneurysms. *J Vasc Surg* 2002; 36: 1260–1263.
14. Piffaretti G, Tozzi M, Lomazzi C, et al. Splenic artery aneurysms: postembolization syndrome and surgical complications. *Am J Surg* 2007; 193: 166–170.
15. Hashizume M, Ohta M, Ueno K, et al. Laparoscopic ligation of splenic artery aneurysm. *Surgery* 1993; 113: 352–354.
16. Pietrabissa A, Ferrari M, Berchiolli R, et al. Laparoscopic treatment of splenic artery aneurysms. *J Vasc Surg* 2009; 50: 275–279.
17. Huang YK, Hsieh HC, Tsai FC, et al. Visceral artery aneurysm: risk factor analysis and therapeutic opinion. *Eur J Vasc Endovasc Surg* 2007; 33: 293–301.
18. Pasha SF, Gloviczki P, Stanson AW, Kamath PS. Splanchnic artery aneurysms. *Mayo Clin Proc* 2007; 82: 472–479.
19. O'Driscoll D, Olliff SP, Olliff JF. Hepatic artery aneurysm. *Br J Radiol* 1999; 72: 1018–1025.
20. Abbas MA, Fowl RJ, Stone WM, et al. Hepatic artery aneurysm: factors that predict complications. *J Vasc Surg* 2003; 38: 41–45.
21. Salam TA, Lumsden AB, Martin LG, Smith RB III. Nonoperative management of visceral aneurysms and pseudoaneurysms. *Am J Surg* 1992; 164: 215–219.
22. Pilleul F, Beuf O. Diagnosis of splanchnic artery aneurysms and pseudoaneurysms, with special reference to contrast enhanced 3D magnetic resonance angiography: a review. *Acta Radiol* 2004; 45: 702–708.
23. Upchurch GR Jr, Zelenock GB, Stanley JC. Splanchnic artery aneurysms. In: Rutherford RB, ed. *Vascular Surgery*. 6th ed. Philadephia, PA: Saunders; 2005; 1565–1581.
24. Stone WM, Abbas M, Cherry KJ, et al. Superior mesenteric artery aneurysms: is presence an indication for intervention? *J Vasc Surg* 2002; 36: 234–237.
25. Stone WM, Abbas MA, Gloviczki P, et al. Celiac arterial aneurysms: a critical reappraisal of a rare entity. *Arch Surg* 2002; 137: 670–674.
26. Grotemeyer D, Duran M, Park EJ, et al. Visceral artery aneurysms–follow-up of 23 patients with 31 aneurysms after surgical or interventional therapy. *Langenbecks Arch Surg* 2009; 394: 1093–1100.
27. Waldenberger P, Bendix N, Petersen J, et al. Clinical outcome of endovascular therapeutic occlusion of the celiac artery. *J Vasc Surg* 2007; 46: 655–661.
28. Leon LR, Jr, Mills JL Sr, Jordan W, et al. The risks of celiac artery coverage during endoluminal repair of thoracic and thoracoabdominal aortic aneurysms. *Vasc Endovascular Surg* 2009; 43: 51–60.
29. Vaddineni SK, Taylor SM, Patterson MA, Jordan WD Jr. Outcome after celiac artery coverage during endovascular thoracic aortic aneurysm repair: preliminary results. *J Vasc Surg* 2007; 45: 467–471.
30. Libicher M, Reichert V, Aleksic M, et al. Balloon occlusion of the celiac artery: a test for evaluation of collateral circulation prior endovascular coverage. *Eur J Vasc Endovasc Surg* 2008; 36: 303–305.
31. Dinter DJ, Rexin M, Kaehler G, Neff W. Fatal coil migration into the stomach 10 years after endovascular celiac aneurysm repair. *J Vasc Interv Radiol* 2007; 18(pt 1): 117–120.
32. Sessa C, Tinelli G, Porcu P, et al. Treatment of visceral artery aneurysms: description of a retrospective series of 42 aneurysms in 34 patients. *Ann Vasc Surg* 2004; 18: 695–703.
33. Finley DS, Hinojosa MW, Paya M, Imagawa DK. Hepatic artery pseudoaneurysm: a report of seven cases and a review of the literature. *Surg Today* 2005; 35: 543–547.

SPECIAL FEATURE

Toward Person-Centered Medicine: From Disease to Patient to Person

Juan Mezzich, MD, PhD,[1] Jon Snaedal, MD,[2] Chris van Weel, MD, PhD,[3]
and Iona Heath, MB, BChir[4]

[1]World Psychiatric Association President, 2005-2008
[2]World Medical Association President, 2007-2008
[3]World Organization of Family Doctors, 2007-2010
[4]Royal College of General Practitioners President, 2009-2012

Throughout its history, medicine has struggled with the ambiguity between that which human beings hold in common and that which is unique to each particular individual; between that which is shared by all those with a named disease and that which is unique about each person's illness experience. Early traces of the contemporary debate about the human face and soul of medicine can be found in the Sanskrit term for health, *hal*, meaning wholeness, as well as in the Ayurvedic and Chinese medical traditions, ancient but still practiced. These traditions focus on the patient's total health rather than simply on disease, and attempt to articulate a comprehensive and harmonious understanding of health while promoting a highly personalized approach to the treatment of specific diseases and to the enhancement of well-being.[1] Similarly, the roots of the Western medical tradition–Platonic, Aristotelian, and Hippocratic–involve a holistic perspective, according to which if the whole is not well, it is impossible for the part to be well.[2] Remarkably, the modern definition of health enshrined in the World Health Organization's (WHO) 1946 constitution remains congruent with traditional aspirations and refers to a dynamic state of complete physical, emotional, and social well-being, and not only the absence of disease.

However, despite this definition, the attitudes and practice of much of modern medicine have become profoundly disease focused and organ specific with ever-increasing specialization. The benefits of scientific progress on disease and innovative treatment technologies may be squandered and hindered if they are not accompanied by understanding of the patient's concerns and needs.[3] Subspecialization has often led to fragmentation and compartmentalization of care. Ever more pervasive commercialization[4] has tended to render medicine a privilege and a commodity rather than a human right, and has emphasized what can be packaged at the expense of what is subtle or intricate in health care or human development. Impaired and weak communication between clinicians and patients has constrained understanding and collaboration and undermined the effective coordination of health care. A narrow and rigid focus on evidence-based practice has marginalized the values and aspirations of patients seeking help.[5] Effective attention to comorbidity demands care that is person centered[6] and relationship based.[7] The limitations of disease-specific approaches in the context of the growing prevalence of comorbidity are becoming more obvious.[8] In response to these problems and limitations, several clinical developments have started to emerge, including the engagement of biopsychosocial concepts and the integration of health and social services as well as initiatives on fully personalized medicine, on the "recovery" of the whole person, and on the "healing relation" between clinician and patient.[7] As a result, the biographical and social contexts of the patient's condition with attention to personal, family, and carer resources are receiving increasing attention.

Address Correspondence to:

Juan Mezzich
Professor of Psychiatry
Mount Sinai School of Medicine
New York, NY
Email: juanmezzich@aol.com

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20187

© 2010 Mount Sinai School of Medicine

Personalized and contextualized care is becoming a major demand, in line with the dictum of the philosopher José Ortega y Gasset: "I am I and my circumstance." Critically, contextualization encompasses cultural diversity as well as spirituality. Respect for the patient's autonomy, values, and dignity represents a fundamental recognition of his or her personhood and an ethical imperative.[7] To illustrate, respect for the person behind a dementing illness is being considered essential to an unprecedented extent, persuasively prompting person-centered care.[10] Also crucial for clinical care is engaging patients thoughtfully, individually and in groups, including those critical of doctors, so as to reaffirm the dialogic basis of the medical profession.[11] In parallel, there is renewed attention to the person of physicians and other health professionals, involving promotion of their wellness and their cultivation as whole persons well grounded on science and ethically committed to care and to help.

Slowly, these concepts are finding their way into evidence.[12] And it is important to explore this further to strengthen the perspective's scientific base, and in particular to build on the importance of empathy and the physician-patient relationship[13] and of treating patients as worthy ends rather than as means.[3] All these trends seem to argue that a more comprehensive, holistic, and person-centered perspective, combining science and humanism, is needed to enhance and reinforce the foundations of medical care.[14,15]

Global health and medical organizations are beginning to respond to this agenda. The World Medical Association has an enduring concern for the promotion of medical ethics reflected in the development and periodic review of its well-recognized Declaration of Helsinki. It recently published the exemplary book of profiles *Caring Physicians of the World*,[16] and its leaders are often speaking on the obligation to respect human values.[17] The World Organization of Family Doctors (Wonca) and its national counterparts have a well-established commitment to the holistic care of patients in the context of family and community.[7] Family physicians have been recently expanding their professional approach from patient centered to person centered, perceiving the latter as more effective and relevant to clinical responsibilities that include prevention and health promotion.[18]

Nowhere has the tension between a reductionistic focus on disease and the consideration of the whole person been more apparent than in the mental health field. The promotion of a psychiatry *of* the person (of the totality of the person's health, including both ill and positive aspects), *for* the person (to assist in the fulfilment of each person's health aspirations and life project, and not merely disease management), *by* the person (with clinicians extending themselves as total human beings with high ethical aspirations), and *with* the person (in respectful collaboration and in an empowering manner) summarizes a valuable response from the World Psychiatric Association through an initiative on Psychiatry for the Person,[14] recently endorsed by the World Federation for Mental Health with its professional, patient, family, and advocate constituents.

To bolster this process, the above-mentioned organizations started collaboration with the World Federation of Neurology, the World Federation for Medical Education, the Council for International Organizations of Medical Sciences, the International Council of Nurses, and the International Alliance of Patients' Organizations, among other bodies. This resulted in the first Geneva Conference on Person-Centered Medicine, in May 2008,[19] followed by the Second Geneva Conference in May 2009[20] and the Third Geneva Conference scheduled for May 2010 with an even larger number of cooperating international bodies. Among these, of particular note, has been the WHO, which, on the basis of the latest World Health Assembly's[21] strong endorsement of people-centered care, has become a major co-organizer of the Geneva conferences. Emerging from this process and now coordinating it and undertaking new initiatives is the International Network for Person-Centered Medicine.[22]

We call upon other scientific societies, academic institutions, and public organizations in the medical and health care field for support and collaboration in building a person-centered medicine dedicated to the promotion of health as a state of physical, mental, social, and spiritual well-being as well as to the reduction of disease, and founded on mutual respect for the dignity and responsibility of each individual person. We are keen to extend the focus of medicine from disease to patient to person by articulating science and humanism and gaining better insight into how the different dimensions of the person-centered approach affect the process and outcome of care, and working respectfully with our patients toward every person's greater health and life's fulfillment.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Patwardhan B, Warude D, Pushpangadan P, Bhatt N. Ayurveda and traditional Chinese medicine: a

comparative overview. *Evid Based Complement Alternat Med* 2005; 2: 465–473.

2. Christodoulou GN, ed. *Psychosomatic Medicine*. New York, NY: Plenum Press; 1987.
3. Heath I. Promotion of disease and corrosion of medicine. *Can Fam Physician* 2005; 51: 1320–1322.
4. World Health Organization. *Primary Care, Now More than Ever–The World Health Report 2008*. Geneva, Switzerland: World Health Organization; 2008.
5. Miles A. Loughlin M, Polychronis A. Evidence-based healthcare, clinical knowledge and the rise of personalised medicine. *J Eval Clin Pract* 2008; 14: 621–649.
6. Salloum IM, Mezzich JE, Cordoba SE. Conceptual and methodological considerations in comorbidity. In: Salloum IM, Mezzich JE, eds. *Psychiatric Diagnosis: Challenges and Prospects*. Chichester, UK; Wiley-Blackwell; 2009 p. 189–196.
7. McWhinney IR. *A Textbook of Family Medicine*. 1st ed. Oxford: Oxford University Press; 1989.
8. Van Weel C, Schellevis FG. Comorbidity and guidelines: conflicting interests. *Lancet* 2006; 367: 550–551.
9. Fulford KWM, Dickenson D, Murray TH, eds. *Healthcare Ethics and Human Values. An Introductory Text with Readings and Case Studies*. Malden, MA: Wiley-Blackwell; 2002.
10. Kitwood T. The uniqueness of persons in dementia. In: Marshall M. ed. *The State of the Art in Dementia Care*. London, UK; Centre for Policy on Ageing Publications; 1995.
11. Mezzich JE. The dialogal basis of our profession: psychiatry with the person. *World Psychiatry* 2007; 6: 129–130.
12. Di Blasi Z, Harkness E, Ernst E, et al. Influence of context effects on health outcomes: a systematic review. *Lancet* 2001; 357: 757–762.
13. Van Os TW, van den Brink RH, Tiemens BG, et al. Communicative skills of general practitioners augment the effectiveness of guideline-based depression treatment. *J Affect Disord* 2005; 84: 43–51.
14. Mezzich JE. Psychiatry for the Person: articulating medicine's science and humanism. *World Psychiatry* 2007; 6: 65–67.
15. Miles A. On a medicine of the whole person: away from scientistic reductionism and towards the embrace of the complex in clinical practice. *J Eval Clin Pract* 2009; 15: 941–949.
16. Coble Y. *Caring Physicians of the World*. Ferney-Voltaire, France: World Medical Association; 2005.
17. Snaedal J. Presidential address. *World Med J* 2007; 53: 101–102.
18. Epperly T. On the person-centered medical home. Paper presented at: Second Geneva Conference on Person-Centered Medicine; May 28–29, 2009; Geneva, Switzerland.
19. Mezzich JE, Snaedal J, Van Weel C, Heath I. Introduction to conceptual explorations on person-centered medicine. *Int J Integr Care* 2010; 10: e022.
20. Mezzich JE. The Second Geneva Conference on Person-Centered Medicine. *World Med J* 2009; 55: 100–101.
21. World Health Organization. Sixty-second World Health Assembly resolutions. Geneva, Switzerland: World Health Organization; 2009.
22. Mezzich JE, Snaedal J, van Weel C, Heath I. The International Network for Person-centered Medicine: background and first steps. *World Med J* 2009; 55: 104–107.

SPECIAL FEATURE

Sports-Related Injuries in the Young Female Athlete

Alexis C. Colvin, MD, and Abigail Lynn, MD

Mount Sinai School of Medicine, New York, NY

OUTLINE

Soccer
Gymnastics
Volleyball
Dance
Cheerleading
Discussion

ABSTRACT

The number of females participating in sports continues to increase. Adolescent and preadolescent females are at a risk injury to both their open growth plates as well as their joints. The purpose of this article is to review the common injuries seen with the most popular sports with females. ***Mt Sinai J Med 77:307–314, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: adolescent females, female athletes, sports injuries.

Female participation in sports is increasing at an exponential rate. In a 2002 survey, girls comprised 44% of all organized sports team members for children aged 6 to 17.[1] Between 1971 and 2001, there was an 800% increase in the number of females participating in high school sports.[1] Furthermore, a 2008 survey of female participation in intercollegiate sports found the highest-ever number of women's teams per school.[2] Although there are obvious benefits to participation in sports (health, teamwork, etc.), young females are also predisposed to certain injuries.

Address Correspondence to:
Alexis C. Colvin
Mount Sinai School of Medicine
New York, NY
Email: alexis.colvin@mountsinai.org

In young athletes, the growth plates, also known as the epiphyseal plates or physes, are at risk for injury while they remain open. Furthermore, skeletally immature athletes are also vulnerable to injury due to an imbalance between strength and flexibility.[3] Between the ages of 6 and 14, limb mass increases twice as much as limb length.[4] This may also lead to a possible imbalance because muscles need to generate a larger force to move limbs.[4] Also, although there is an increase in bone stiffness, resistance to impact decreases and a sudden overload may cause bones to bend or buckle.[3]

The purpose of this article is to review common sports played by young females, common injuries sustained, and prevention of those injuries.

SOCCER

Soccer is one of the most popular sports worldwide. Female participation is also rapidly increasing. There was a 20% increase in the number of females age 6 and older participating in soccer from 1987 to 1999.[1]

With the increase in participation has also come an increase in injuries. This is reflected in several statistics. First, there was an increase in the number of absolute soccer-related injuries in pediatric girls presenting to emergency departments from 1990 to 2003.[5] The injury rate for girls also increased from 1.14 to 1.63 per 1000 population.[5] Girls were more likely than boys to have ankle and knee injuries (typically sprains or strains).[5] When insurance claims from youth soccer leagues over a 5-year period were examined, females demonstrated a higher ratio of knee injuries to all injuries and a higher ratio of anterior cruciate ligament (ACL) injuries to all injuries.[6] Swenson *et al.* (2009) looked at high school soccer athletes from 2005 to 2008 and found that girls had higher rates of recurrent injuries than boys, with recurrent injuries most often affecting the ankle (28.3%) and knee (16.8%), usually as ligament sprains

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20179

© 2010 Mount Sinai School of Medicine

(34.9%) and muscle strains (13.3%).[7] ACL injuries represented 37% of total knee injuries from females versus 24% from males.[6] In fact, females who play high school and collegiate soccer year-round have a 5% ACL tear rate.[8] Furthermore, female soccer players are nearly 3 times more likely than male soccer players to tear their ACL.[8]

The ACL prevents anterior translation of the tibia relative to the femur. Postinjury, patients typically recall hearing a "pop," experiencing immediate swelling, and being unable to continue the activity. The Lachman and anterior drawer tests can be used to assess an ACL rupture on physical exam. Radiographs are typically negative, whereas an MRI can confirm the diagnosis as well as indicate any other pathology, such as damage to the meniscus, cartilage, or other ligaments. ACL reconstruction is typically recommended for young patients to help prevent further meniscal injury and allow return to activity.

Several factors for the increased risk of ACL injury among females have been proposed, including differences in anatomy, hormone levels, and neuromuscular factors.[9] However, there is no consensus on differences in pelvis shape and/or lower-extremity alignment being responsible for increased injury rates.[10,11] There is also no agreement as to whether femoral anatomy or ACL size can predispose females to potential injury.[11–15] Conflicting evidence also exists on the effect of estrogen on ACL injury.[16,17] Currently, neuromuscular factors seem to be the most important contributors to the increased risk of ACL injury in females and may influence differential ACL injury rates between males and females.[9] Several authors have found a decreased incidence of ACL injury with targeted training programs.[8,18–20]

In male teenage soccer players, Junge *et al.* (2002) found that an injury-prevention program helped to reduce the number of soccer injuries.[20] This program focused on educating both coaches and players on injury prevention through improving flexibility and trunk, knee, and ankle strength. The authors found that this program was more effective in lower-skilled players.

Prodromos *et al.* (2007) performed a meta-analysis and found that injury-reduction training helped reduced ACL tear incidence in female soccer players.[8] Other authors have had similar findings, with benefits to females involved in other sports as well.[18,19]

GYMNASTICS

Elite gymnasts can begin as early as age 4 and spend up to 50 hours per week practicing and competing.[21] Not surprisingly, gymnastics has one of the highest injury rates of all of girls' sports.[22] Participation in gymnastics is associated with an increased risk of stress fractures in females younger than 17 years.[23] Furthermore, gymnastics also produces the highest number of injuries requiring surgeries.[24] Girls aged 12 to 17 have a greater number of injuries sustained (7.4 per 1000 participants) than girls aged 6 to 11 (3.6 per 1000 participants).[22] The potential areas for injury include the wrist and hand, the spine, and the ankle and foot.

Wrist pain is a very common complaint with gymnasts. Unlike other sports, gymnastics uses the upper extremity for weight bearing.[22,25] The wrist is also subject to multiple stresses in different directions.[25] In a survey of nonelite, primarily female gymnasts of an average age of 11.8 years, 73% had wrist pain.[26] Gymnasts with wrist pain were older, trained more hours per week at a higher skill level, and began gymnastics at an older age (7.0 years versus 5.1 years).[26] The age discrepancy is most likely due to a period of relative weakness of the bone during the adolescent growth spurt.[27] During peak velocity of growth, there is an imbalance between bone growth and bone mineralization.[27] When gymnasts present with wrist pain, it is typically located over the dorsal wrist and provoked with floor-based exercise and activities.[26]

The differential diagnoses of wrist pain can be divided based on location.[21] The most common diagnosis is a distal radius stress fracture.[25] Other radial-sided wrist injuries include scaphoid impaction syndrome, scaphoid fractures, and scaphoid stress reactions or stress fractures.[21] Ulnar-sided injuries include avascular necrosis of the capitate, triangular fibrocartilage complex (TFCC) tears, ulnar impaction syndrome, and lunotriquetral impingement.[21,25] Dorsal injuries include dorsal impingement, ganglia, and carpal instability.[21]

In addition to constant overuse, another potential risk factor for wrist pain is the gymnastics equipment itself.[28] Very soft mats may exaggerate dorsiflexion of the wrist.[28] Performance of a twisting vault leads to hyperextension of the wrist.[28] Tumbling on the balance beam causes torsion on the forearms with pronation and supination.[28]

Regarding the prevention of wrist pain, there have been no clinical studies on skeletally immature female gymnasts and the use of wrist guards. However, 2 cadaver studies have shown that the use of wrist braces and guards may help by sharing load and absorbing impact energy.[29,30]

Other preventive measures include gradual increases in training loads and reduction of training during periods of growth.[26,31] Having gymnasts alternate swinging and support-type movements may also help to reduce the intensity of compressive load on the wrist.[32] As with all injuries, early detection and treatment can help to limit the extent of the injury and hasten recovery.[33]

The spine is also at increased risk for injury in gymnasts. The different body positions and landings result in a varied amount of hyperflexion and hyperextension of the spine.[34] An increased number of practice and competition hours, more than 16 per week, is associated with a significant increase in back injuries in female gymnasts aged 12 to 26 years ($P < 0.0002$).[35] Findings that have been seen on radiographs and MRIs of female gymnasts have included degenerated and herniated intervertebral discs, thoracic compression fractures, kyphosis, retrolisthesis, and Schmorl's node.[36]

Gymnasts with injuries related to the spine can present with injuries in the cervical, thoracic, or lumbar area. In a review of spine injuries in gymnasts from 1954 to 1984, 28 of 38 spinal injuries occurred in the cervical spine.[37] These injuries occurred secondary to the gymnasts landing on their hands, resulting in the force being transmitted to the cervical spine.[37] The primary cause of injury was failure of supervision.[37]

Spondylolysis, a fracture of the pars interarticularis, results from repetitive hyperextension and rotation of the lumbar spine.[34] Gymnasts typically present with dull low back pain exacerbated by certain maneuvers that cause hyperextension, hyperextension and axial load, increased axial load, or rotation of the lumbar spine.[38] A provocative test that can elicit the symptoms is the "one-legged hyperextension" test, performed with the patient in a standing position.[38] The neurologic exam is normal.[39] Although oblique radiographs can diagnostic,[40] some authors have recommended only taking AP and lateral views to limit the amount of radiation to the gonads.[34] Currently, a single-photon emission computed topography (SPECT) scan is recommended for imaging.[34] With a positive SPECT but negative CT, the recommended treatment is rest for at least 12 weeks or until symptoms resolve.[41] Other treatments that can be considered are bracing and physical therapy, with a focus on spine-stabilization exercises and sport-specific retraining.[41] The average return to pain-free competition is 7.3 months.[38]

Spondylolisthesis is defined as the slippage of one vertebra on another, and the movement can be forward (anterolisthesis) or backward (retrolisthesis).[42] Gymnasts are more likely to develop isthmic spondylolisthesis, seen as an anterolisthesis resulting from bilateral pars interarticularis fractures.[34] The preadolescent incidence of spondylolisthesis is 4.4%, with a slippage of as much as 28%.[43] After adolescence, it is unusual to have progression of the slippage.[42]

If the slippage is 25% to 50% in an asymptomatic child, or <50% in a symptomatic child, consideration should be given to limiting gymnastics activities.[42] A child with a slippage >50% or with persistence of symptoms with a 25% to 50% slippage should be evaluated for a possible fusion.[42]

Athletes with spondylolysis may benefit more from a nonoperative protocol than those with spondylolisthesis.[40] Blanda *et al.* (1993) compared their results with a treatment protocol involving no sports and full-time use of a lumbosacral orthosis until the athlete was able to perform lumbar extensions.[40] Once the athlete was pain-free, hamstring stretching, pelvic tilts, and abdominal strengthening exercises were continued for a minimum of 6 months.[40] The authors found greater success in children with spondylolysis (82%) versus spondylolisthesis (40%).[40] They concluded that surgical treatment would be more likely for those athletes with a longer history of symptoms.[40]

The foot and ankle are also susceptible to injury in gymnasts. In a 3-year review of 14 elite female gymnasts (aged 14 to 21 years) who presented with foot or ankle injuries, all patients required surgery.[24] There were 2 skeletally immature gymnasts, 1 who was unable to return after sustaining osteochondral talus injury.[24] The other skeletally immature gymnasts needed later surgery to correct malalignment after injury to the growth plate.[24] The injuries sustained included 5 Lisfranc fracture-dislocations, 5 talocalcaneal fractures, 2 multiple metatarsal fractures, 1 medial malleolar fracture, 1 phalangeal fracture, and 1 sesamoid fracture (Figure 1).[24] Six of the 14 injuries occurred during dismount from the vault.[24] Indeed, loads from landing are most substantial at the tibiotalar and talonavicular joint and can reach >23 × body weight.[44]

VOLLEYBALL

Volleyball has become an increasingly popular sport among females. Because of the significant amount of jumping involved, many of the injuries sustained by young athletes in volleyball are similar to those sustained in soccer and basketball. However, volleyball players may actually have a lower rate of ACL tears.[8,45] This may be partly attributed to female volleyball players landing more symmetrically when performing a jump-drop landing versus female basketball players, who tend to land more

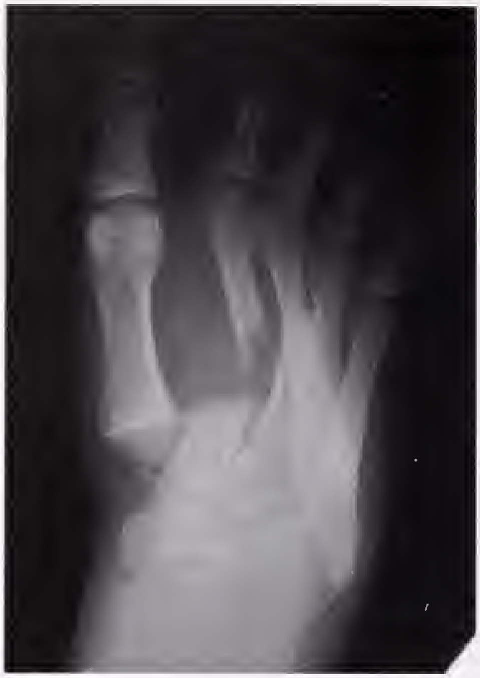


Fig 1. Anteroposterior radiograph of Lisfranc injury in a gymnast.

asymmetrically.[46] Females in general have different mechanics when it comes to landing compared with males. Females do not flex their hips and knees as far as their male counterparts. In addition, females have a higher ground reaction force.[47] Those female athletes who do possess increased dynamic valgus and high abduction loads are the ones who are at increased risk of ACL injury. In a prospective study done in 205 female athletes in soccer, basketball, and volleyball, it was found that those who sustained ACL tears have decreased neuromuscular control and increased valgus joint loading. In fact, the knee abduction moment was both highly specific and sensitive in predicting those females who ended up sustaining an ACL tear. Knee abduction angle, knee abduction moment, and ground reaction force were all greater in those who had ACL injuries.[48] As has been previously stated, jump-training programs have been developed to strengthen and correct these mechanics of landing. Although some studies have shown that these programs do decrease the rate of ACL tears, this is still up for debate. One prospective cohort study with nearly 1500 female high school athletes playing soccer, basketball, and volleyball showed no significant difference in using a plyometric-based training program in reducing the incidence of noncontact ACL injuries.[49]

Even if these young, athletic volleyball players jump and land with better mechanics, they are still at risk for other injuries. Repetitive jumping can cause inflammation of the patellar tendon, leading to "jumper's knee," or patellar tendonitis. In the growing child, this pulling of the tendon can cause repetitive microtrauma at its attachments to the growing bone. If the pulling is on the insertion of the tibial tubercle, then the disease is termed Osgood-Schlatter. If the pulling is on the origin of the distal pole of the patella, then the disease is termed Sindig-Larsen-Johanssen. Jumper's knee is common in elite young volleyball players and is usually diagnosed clinically, but it can also be diagnosed by the presence of neovascularization seen on ultrasound.[50] It usually responds to rest, ice, and anti-inflammatories.

The most common injury in volleyball players is an ankle sprain. Wearing bilateral double upright padded ankle braces appears to decrease the incidence of this relatively common injury.[51]

Volleyball players are also at risk for shoulder injuries due to the frequent overhead activity. It has been hypothesized that these injuries may be due to an imbalance between the internal and external rotators of the shoulder.[52]

DANCE

Dance is highly demanding and takes dedication to long hours of training. Ninety percent of dancers will sustain an injury at some point in their lifetime as a dancer. Force on the forefoot is extraordinary, especially in the ballet en pointe position.[53] Lower-extremity injuries such as sprains and tendonitis are common in the dancer.[54]

A dancer may have different anatomy and kinematics compared with a nondancer. One study has suggested a change in the actual bony geometry of young classical ballet dancers who train at a higher intensity. Those who trained for more than 6 hours a day had significantly less femoral torsion.[55] This change in bony anatomy may predispose ballet dancers to osteoarthritis. Young dancers who work professionally for more than 10 years have an increased incidence of arthritis in the knees, ankles, and first metatarsal by their mid 20s to mid 40s compared with nondancers.[56]

During the course of a young dancer's career, she may end up developing foot disorders, including a symptomatic os trigonum, tears of the flexor hallucis longus, or a compression neuropathy, including tarsal tunnel syndrome, medial hallucal nerve compression, anterior tarsal tunnel syndrome, superficial and deep peroneal nerve entrapment, and sural nerve entrapment.[57] Os trigonum is the term for an accessory bone that can impinge and cause pain, swelling, and neurologic changes. If the patient fails nonsurgical management, it can respond favorably to simple excision of the os.[58] Dancers may also get tears in the flexor hallucis longus tendon in their foot. These also respond well to surgical treatment.[59]

Ankle inversion injuries are ubiquitous in dancers. Strengthening both the static and dynamic stabilizers is of paramount importance in dancers due to the extreme stress placed on the ankle.[60] Once a dancer sustains a sprain, she is at a higher risk for sustaining a sprain in the contralateral ankle.[61] Recurrent ankle sprains may lead to anterior impingement of the soft tissue or bony structures. This impingement causes pain and can preclude dancing. In 1 retrospective case series of 11 dancers, all eventually returned to dance after arthroscopic bony and soft-tissue debridement of the offending structure.[62]

Stress fractures can also occur in dancers, specifically in the fibula, calcaneus, navicular, and second metatarsal, as opposed to the tibia, in which stress fractures are more common in distance and track runners.[63,64] Prolonged healing times associated with these stress fractures can be a frustrating problem for the dancer. In refractory cases of tibial stress fractures, surgical treatment such as intramedullary nailing or drilling and bone grafting allows for healing of the fracture, resolution of symptoms, and return to dance.[65]

Dancers have a lower incidence of ACL tears relative to other sports, which may be explained by the fact that dance training requires practicing jump technique, similar to the neuromuscular training programs mentioned previously.[66] In fact, there is no gender disparity in the incidence of ACL tears in dancers. Kinematics and ground reaction forces during single-leg jump landings showed no difference between genders.[67] Dancers typically sustain an ACL tear by landing on the leg in a turned-out position with a valgus force at the knee. Interestingly, in 1 retrospective study of 253 dancers, all 6 dancers who sustained an ACL tear had the tear occur on the left side.[68]

Surgery can often be career ending for a dancer.[69] Thus, prevention of injuries is critical. In dancers, injuries may be secondary to a "kinetic chain dysfunction," or an abnormal motion causing the injury. If the dancer corrects this abnormal motion, and not simply the resultant injury, it could potentially prevent future injuries.[70] For instance, if a dancer has a specific tendon injury, it is important to address the overall malalignment which could have predisposed her to that injury.[71] Having a multidisciplinary team that is familiar with dance mechanics can also help with early detection of poor mechanics and thus prevention of injury.[72]

CHEERLEADING

The number of cheerleading-related visits to emergency rooms has doubled over the past decade.[73] Furthermore, cheerleading has the greatest number of catastrophic injuries.[74] When compared with all other sports in which females participate, cheerleading accounts for more than half of all catastrophic injuries.[75] Catastrophic injuries are rare, but tragic, events. They include head injury, skull fractures, cervical spine injury, heat stroke, commotio cordis, heat illness, and exertional hyponatremia.[77,76]

Participation in sports does have a plethora of positive effects on the young teenager–physiologically, psychologically, and socially. Even so, one should also not assume that those involved in athletics are less likely to be involved with alcohol or other substance use. In fact, females who are in out-of-school programs for dance, cheerleading, or gymnastics are at an increased risk of using at least 1 substance. Interestingly, females who participate in these same 3 sports organized within the school are at a decreased risk for substance use.[77]

DISCUSSION

With the increasing number of females participating in sports, there is the potential for more injuries. Prevention of injuries in young females should be a priority. Once a female athlete in basketball, soccer, or volleyball has an injury, she is actually much more likely to sustain a new injury than a reinjury.[78]

Prevention begins with the preparticipation history and physical exam to determine any predisposing conditions to injury or life-threatening conditions.[79] Appropriate medical coverage at sporting events should also be ensured.[80] Proper coaching, equipment, and surface playing conditions, as well as proper officiating and adequate hydration of the athletes, are also important.[80]

The team physician also plays an important role in helping to screen for the female athlete triad. The triad is defined as menstrual irregularity, disordered eating (leading to amenorrhea), and premature osteoporosis.[81] Disordered eating actually does not independently directly correlate with getting a stress fracture; however, being active more than 16 hours per week is associated with disordered eating.[23] Athletes who have had amenorrhea for >6 months have lower bone mineral density (BMD) than control subjects.[82] These females have an increased risk of stress fractures.[83] Athletes who bear weight using their upper extremities tend to get stress fractures distal to the elbow; athletes who swing will tend to get them in their lower ribs; throwing athletes can get them in the shoulder girdle area; and weight lifters can get them throughout the upper extremity.[84]

The treatment should be multidisciplinary, involving parents, the coach, and the athletic trainer.[80] Future research should focus on prevention and treatment of injuries in female athletes as most of our current knowledge of musculoskeletal injuries is derived from studies of male athletes.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Women's Sports Foundation. Women's Sports & Physical Activity Facts & Statistics. http://www.womenssportsfoundation.org/binary-data/WSF_ARTICLE/pdf_file/28.pdf. Updated December 12, 2007.
2. Acosta RV, Carpenter LJ. Women in intercollegiate sport: a longitudinal, national study. Thirty three year update, 1977–2010. http://www.acostacarpenter.org.
3. Maffulli N, Baxter-Jones AD. Common skeletal injuries in young athletes. *Sports Med* 1995; 19: 137–149.
4. Hawkins D, Metheny J. Overuse injuries in youth sports: biomechanical considerations. *Med Sci Sports Exerc* 2001; 33: 1701–1707.
5. Leininger RE, Knox CL, Comstock RD. Epidemiology of 1.6 million pediatric soccer-related injuries presenting to US emergency departments from 1990 to 2003. *Am J Sports Med* 2007; 35: 288–293.
6. Shea KG, Pfeiffer R, Wang JH, et al. Anterior cruciate ligament injury in pediatric and adolescent soccer players: an analysis of insurance data. *J Pediatr Orthop* 2004; 24: 623–628.
7. Swenson DM, Yard EE, Fields SK, Comstock RD. Patterns of recurrent injuries among US high school athletes, 2005–2008. *Am J Sports Med* 2009; 37: 1586–1593.
8. Prodromos CC, Han Y, Rogowski J, et al. A meta-analysis of the incidence of anterior cruciate ligament tears as a function of gender, sports, and a knee injury-reduction regimen. *Arthroscopy* 2007; 23: 1320–1325.
9. Griffin LY, Agel J, Albohm MJ, et al. Noncontact anterior cruciate ligament injuries: risk factors and prevention strategies. *J Am Acad Orthop Surg* 2000; 8: 141–150.
10. Haycock CE, Gillette JV. Susceptibility of women athletes to injury. *JAMA* 1976; 236: 163–165.
11. Zelisko JA, Noble HB, Porter M. A comparison of men's and women's professional basketball injuries. *Am J Sports Med* 1982; 10: 297–299.
12. Emerson RJ. Basketball knee injuries and the anterior cruciate ligament. *Clin Sports Med* 1993; 12: 317–328.
13. Gray J, Taunton JE, McKenzie DC, et al. A survey of injuries to the anterior cruciate ligament of the knee in female basketball players. *Int J Sports Med* 1985; 6: 314–316.
14. Hutchinson MR, Ireland ML. Knee injuries in female athletes. *Sports Med* 1995; 19: 288–302.
15. Shelbourne KD, Davis TJ, Klootwyk TE. The relationship between intercondylar notch width of the femur and the incidence of anterior cruciate ligament tears. A prospective study. *Am J Sports Med* 1998; 26: 402–408.
16. Myklebust G, Maehlum S, Holm I, Bahr R. A prospective cohort study of anterior cruciate ligament injuries in elite Norwegian team handball. *Scand J Med Sci Sports* 1998; 8: 149–153.
17. Wojtys EM, Huston LJ, Lindenfeld TN, et al. Association between the menstrual cycle and anterior cruciate ligament injuries in female athletes. *Am J Sports Med* 1998; 26: 614–619.
18. Caraffa A, Cerulli G, Projetti M, et al. Prevention of anterior cruciate ligament injuries in soccer: a prospective controlled study of proprioceptive training. *Knee Surg Sports Traumatol Arthrosc* 1996; 4: 19–21.
19. Hewett TE, Lindenfeld TN, Riccobene JV, Noyes FR. The effect of neuromuscular training on the incidence of knee injury in female athletes. A prospective study. *Am J Sports Med* 1999; 27: 699–706.
20. Junge A, Rosch D, Peterson L, et al. Prevention of soccer injuries: a prospective intervention study in youth amateur players. *Am J Sports Med* 2002; 30: 652–659.
21. Gabel GT. Gymnastic wrist injuries. *Clin Sports Med* 1998; 17: 611–621.
22. Singh S, Smith GA, Fields SK, McKenzie LB. Gymnastics-related injuries to children treated in emergency departments in the United States, 1990–2005. *Pediatrics* 2008; 121: 954–960.
23. Loud KJ, Gordon CM, Micheli LJ, Field AE. Correlates of stress fractures among preadolescent and adolescent girls. *Pediatrics* 2005; 115: 399–406.
24. Chilvers M, Donahue M, Nassar L, Manoli A II. Foot and ankle injuries in elite female gymnasts. *Foot Ankle Int* 2007; 28: 214–218.
25. Webb BG, Rettig LA. Gymnastic wrist injuries. *Curr Sports Med Rep* 2008; 7: 289–295.
26. DiFiori JP, Puffer JC, Mandelbaum BR, Mar S. Factors associated with wrist pain in the young gymnast. *Am J Sports Med* 1996; 24: 9–14.
27. Bailey DA, Wedge JH, McCulloch RG, et al. Epidemiology of fractures of the distal end of the radius in children as associated with growth. *J Bone Joint Surg* 1989; 71: 1225–1231.
28. Roy S, Caine D, Singer KM. Stress changes of the distal radial epiphysis in young gymnasts: a report of twenty-one cases and a review of the literature. *Am J Sports Med* 1985; 13: 301–308.
29. Moore MS, Popovic NA, Daniel JN, et al. The effect of a wrist brace on injury patterns in experimentally produced distal radial fractures in a cadaveric model. *Am J Sports Med* 1997; 25: 394–410.
30. Staebler MP, Moore DC, Akelman E, et al. The effect of wrist guards on bone strain in the distal forearm. *Am J Sports Med* 1999; 27: 500–506.
31. Caine D, Cochrane B, Caine C, Zemper E. An epidemiologic investigation of injuries affecting young competitive female gymnasts. *Am J Sports Med* 1989; 17: 811–820.
32. Sands WA. Injury prevention in women's gymnastics. *Sports Med* 2000; 30: 359–373.

33. DiFiori JP, Caine DJ, Malina RM. Wrist pain, distal radial physeal injury, and ulnar variance in the young gymnast. *Am J Sports Med* 2006; 34: 840–849.
34. Kruse D, Lemmen B. Spine injuries in the sports of gymnastics. *Curr Sports Med Rep* 2009; 8: 20–28.
35. Goldstein JD, Berger PE, Windler GE, Jackson DW. Spine injuries in gymnasts and swimmers. An epidemiologic investigation. *Am J Sports Med* 1991; 19: 463–468.
36. Katz DA, Scerpella TA. Anterior and middle column thoracolumbar spine injuries in young female gymnasts. Report of seven cases and review of the literature. *Am J Sports Med* 2003; 31: 611–616.
37. Silver JR, Silver DD, Godfrey JJ. Injuries of the spine sustained during gymnastic activities. *Br Med J (Clin Res Ed)* 1986; 293: 861–863.
38. Jackson DW, Wiltse LL, Dingeman RD, Hayes M. Stress reactions involving the pars interarticularis in young athletes. *Am J Sports Med* 1981; 9: 304–312.
39. Standaert CJ, Herring SA. Spondylolysis: a critical review. *Br J Sports Med* 2000; 34: 415–422.
40. Blanda J, Bethem D, Moats W, Lew M. Defects of pars interarticularis in athletes: a protocol for nonoperative treatment. *J Spinal Disord* 1993; 6: 406–411.
41. Standaert CJ. New strategies in the management of low back injuries in gymnasts. *Curr Sports Med Rep* 2002; 1: 293–300.
42. Lonstein JE. Spondylolisthesis in children: cause, natural history, and management. *Spine* 1999; 24: 2640–2648.
43. Fredrickson BE, Baker D, McHolick WJ, et al. The natural history of spondylolysis and spondylolisthesis. *J Bone Joint Surg* 1984; 66: 699–707.
44. Bruggemann GP. Mechanical load in artistic gymnastics and its relation to apparatus and performance. In: Leglise M, ed. *Symposium Medico-Technique*. Lyss, Switzerland: International Gymnastics Federation; 1999:; 17–27.
45. Iwamoto J, Takeda T, Sato Y, Matsumoto H. Retrospective case evaluation of gender differences in sports injuries in a Japanese sports medicine clinic. *Gend Med* 2008; 5: 405–414.
46. Herrington L. Knee valgus angle during landing task in female volleyball and basketball players. *J Strength Cond Res*; doi: 10.1519/JSC.0b013e3181b62c77.
47. Salci Y, Kentel BB, Heycan C, et al. Comparison of landing maneuvers between male and female college volleyball players. *Clin Biomech* 2004; 19: 622–628.
48. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. *Am J Sports Med* 2005; 33: 492–501.
49. Pfeiffer RP, Shea KG, Roberts D, et al. Lack of effect of a knee ligament injury prevention program on the incidence of noncontact anterior cruciate ligament injury. *J Bone Joint Surg* 2006; 88: 1769–1774.
50. Gisslen K, Gyulai C, Soderman K, Alfredson H. High prevalence of jumper's knee and sonographic changes in Swedish elite junior volleyball players compared to matched controls. *Br J Sports Med* 2005; 39: 298–301.
51. Pedowitz DI, Reddy S, Parekh SG, et al. Prophylactic bracing decreases ankle injuries in collegiate female volleyball players. *Am J Sports Med* 2008; 36: 324–327.
52. Stickley CD, Hetzler RK, Freemyer BG, Kimura IF. Isokinetic peak torque ratios and shoulder injury history in adolescent female volleyball athletes. *J Athl Train* 2008; 43: 571–577.
53. Prisk VR, O'Loughlin PF, Kennedy JG. Forefoot injuries in dancers. *Clin Sports Med* 2008; 27: 305–320.
54. Wiesler ER, Hunter DM, Martin DF, et al. Ankle flexibility and injury patterns in dancers. *Am J Sports Med* 1996; 24: 754–757.
55. Hamilton D, Aronsen P, Loken JH, et al. Dance training intensity at 11–14 years is associated with femoral torsion in classical ballet dancers. *Br J Sports Med* 2006; 40: 299–303.
56. Teitz CC, Kilcoyne RF. Premature osteoarthrosis in professional dancers. *Clin J Sports Med* 1998; 8: 255–259.
57. Kennedy JG, Baxter DE. Nerve disorders in dancers. *Clin Sports Med* 2008; 27: 329–334.
58. Marotta JJ, Micheli LJ. Os trigonum impingement in dancers. *Am J Sports Med* 1992; 20: 533–536.
59. Sammarco GJ, Cooper PS. Flexor hallucis longus tendon injury in dancers and nondancers. *Foot Ankle Int* 1998; 19(6): 356–362.
60. O'Loughlin PF, Hodgkins CW, Kennedy JG. Ankle sprains and instability in dancers. *Clin Sports Med* 2008; 27: 247–262.
61. Hiller CE, Refshauge KM, Herbert RD, Kilbreath SL. Intrinsic predictors of lateral ankle sprain in adolescent dancers: a prospective cohort study. *Clin J Sport Med* 2008; 18: 44–48.
62. Nihal A, Rose DJ, Trepman E. Arthroscopic treatment of anterior ankle impingement syndrome in dancers. *Foot Ankle Int* 2005; 26: 908–912.
63. Goulart M, O'Malley MJ, Hodgkins CW, Charlton TP. Foot and ankle fractures in dancers. *Clin Sports Med* 2008; 27: 295–304.
64. Brukner P, Bradshaw C, Khan KM, et al. Stress fractures: a review of 180 cases. *Clin J Sport Med* 1996; 6: 85–89.
65. Miyamoto RG, Dhotar HS, Rose DJ, Egol K. Surgical treatment of refractory tibial stress fractures in elite dancers: a case series. *Am J Sports Med* 2009; 37: 1150–1154.
66. Liederbach M, Dilgen FE, Rose DJ. Incidence of anterior cruciate ligament injuries among elite ballet and modern dancers: a 5-year prospective study. *Am J Sports Med* 2008; 36: 1779–1788.
67. Orishimo KF, Kremenic IJ, Pappas E, et al. Comparison of landing biomechanics between male and female professional dancers. *Am J Sports Med* 2009; 37: 2187–2193.
68. Meuffels DE, Verhaar JA. Anterior cruciate ligament injury in professional dancers. *Acta Orthop* 2008; 79: 515–518.
69. Brown TD, Micheli LJ. Foot and ankle injuries in dance. *Am J Orthop* 2004; 33: 303–309.
70. Macintyre J, Joy E. Foot and ankle injuries in dance. *Clin Sports Med* 2000; 19: 351–368.
71. Hodgkins CW, Kennedy JG, O'Loughlin PF. Tendon injuries in dance. *Clin Sports Med* 2008; 27: 279–288.
72. Schon LC, Weinfeld SB. Lower extremity musculoskeletal problems in dancers. *Curr Opin Rheumatol* 1996; 8: 130–142.
73. Shields BJ, Smith GA. Cheerleading-related injuries to children 5–18 years of age: United States, 1990–2002. *Pediatrics* 2006; 117: 122–129.

74. Boden BP. Direct catastrophic injury in sports. *J Am Acad Orthop Surg* 2005; 13: 445–454.
75. Jacobson BH, Redus B, Palmer T. An assessment of injuries in college cheerleading: distribution, frequency, and associated factors. *Br J Sports Med* 2005; 39: 237–240.
76. Luckstead EF, Patel DR. Catastrophic pediatric sports injuries. *Pediatr Clin North Am* 2002; 49: 581–591.
77. Moore MJ, Werch CE. Sport and physical activity participation and substance use among adolescents. *J Adolesc Health* 2005; 36: 486–493.
78. Rauh MJ, Macera CA, Ji M, Wiksten DL. Subsequent injury patterns in girls'high school sports. *J Athl Train* 2007; 42: 486–494.
79. Hergenroeder AC. Prevention of sports injuries. *Pediatrics* 1998; 101: 1057–1063.
80. Frank JB, Jarit GJ, Bravman JT, Rosen JE. Lower extremity injuries in the skeletally immature athlete. *J Am Acad Orthop Surg* 2007; 15: 356–366.
81. Teitz CC, Hu SS, Arendt EA. The female athlete: evaluation and treatment of sports-related problems. *J Am Acad Orthop Surg* 1997; 5: 87–96.
82. Myburgh KH, Bachrach LK, Lewis B, et al. Low bone mineral density at axial and appendicular sites in amenorrheic athletes. *Med Sci Sports Exerc* 1993; 25: 1197–1202.
83. Myburgh KH, Hutchins J, Fataar AB, et al. Low bone density is an etiologic factor for stress fractures in athletes. *Ann Intern Med* 1990; 113: 754–759.
84. Sinha AK, Kaeding CC, Wadley GM. Upper extremity stress fractures in athletes: clinical features of 44 cases. *Clin J Sport Med* 1999; 9: 199–202.

NEW DRUGS AND THE FDA

Prasugrel, a New Platelet Inhibitor

Debbie Yee, PharmD, Joanne Meyer, MS, PharmD, Gina Caliendo, PharmD, BCPS, and Veronica Oviedo, PharmD

Mount Sinai Medical Center, New York, NY

Cardiovascular disease is the leading cause of death in America.[1] Acute coronary syndrome (ACS) is a spectrum of cardiovascular diseases that includes unstable angina (UA), non–ST-segment myocardial infarction (NSTEMI), and ST-segment myocardial infarction (STEMI). Percutaneous coronary intervention (PCI) is the recommended treatment for UA, NSTEMI, and STEMI by the American College of Chest Physicians (ACCP).[2,3]

The ACCP guidelines recommend the use of antiplatelet therapy after primary PCI to maintain vessel patency and to improve myocardial perfusion to the infarct area.[2,3] Aspirin 75 to 325 mg daily is recommended for all UA, NSTEMI, and STEMI patients, in addition to a thienopyridine.[2,3] Thienopyridines are adenosine diphosphate (ADP) receptor antagonists that inhibit platelet aggregation and prolong bleeding time.[3] There are currently 3 medications within this class: ticlodipine, clopidogrel, and the novel thienopyridine prasugrel (Effient; Daiichi Sankyo Co., Ltd., Tokyo, Japan; and Eli Lilly and Co., Indianapolis, IN). Ticlodipine is not often used due to frequent gastrointestinal side effects, rash, neutropenia, thrombocytopenia, and liver-function abnormalities.[3] Clopidogrel, however, is widely used because it is better tolerated and has an improved safety profile.[3] Patients should be initiated on clopidogrel with a loading dose of 300 to 600 mg, followed by 75 mg once daily.[2,3]

Like clopidogrel, prasugrel is a prodrug and must be converted to its active metabolite before it can inhibit platelet activation and aggregation.[4,5] The active metabolite binds selectively and irreversibly to ADP P2Y12 receptors and inhibits platelet activation and aggregation for the life of the platelet, which is typically 7 to 10 days.[4] Clinical studies look at the pharmacokinetics, pharmacodynamics, efficacy, and safety of prasugrel in ACS patients undergoing PCI. This article will review the available information to help formulary decision makers decide what is best for their institutions.

Clinical studies have evaluated prasugrel in the treatment of patients with acute coronary syndrome (ACS) undergoing percutaneous coronary intervention (PCI), comparing it with clopidogrel, currently the platelet inhibitor most widely used.

Address Correspondence to:

Debbie Yee
Mount Sinai Medical Center,
New York, NY
Email: debbie.yee@mountsinai.org

PHARMACOKINETICS AND PHARMACODYNAMICS

Prasugrel is converted to an active metabolite by hepatic carboxylesterases followed by a single cytochrome P450 (CYP450)-dependent oxidative step.[4] Conversion is done predominantly through CYP3A4 and CYP2B6, and to a lesser extent CYP2C9 and CYP2C19. Prasugrel is not affected by CYP2C19 genetic variants, and it may be administered with CYP450 inducers or inhibitors.[4] Although prasugrel is a weak inhibitor of CYP2B6, it has not been shown to cause a significant interaction.[4]

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20185

© 2010 Mount Sinai School of Medicine

Prasugrel is associated with a faster onset and greater platelet aggregation inhibition than clopidogrel. This was demonstrated in a randomized, double-blind study of 110 patients with stable coronary artery disease.[7] The mean maximal platelet aggregation was significantly lower 30 minutes after the 60 mg prasugrel loading dose compared with the 600 mg clopidogrel loading dose ($P < 0.001$).[7] Furthermore, P2Y12 receptor blockade was significantly higher after the prasugrel loading dose ($P < 0.001$), and the mean platelet reactivity index was 50% lower after prasugrel maintenance treatment ($P < 0.001$).[7] These results suggest that prasugrel has a rapid onset of action of approximately 30 minutes, and its active metabolite blocks P2Y12 receptors more completely than clopidogrel does.

Prasugrel is associated with a faster onset and greater platelet aggregation inhibition than clopidogrel.

EFFICACY

The Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition with Prasugrel-Thrombolysis in Myocardial Infarction 38 (TRITON-TIMI 38) was a phase III, randomized, double-blind comparative study of ACS patients undergoing PCI.[8] The inclusion criteria for patients with UA/NSTEMI were ischemic symptoms that lasted >10 minutes and occurred within 27 hours before randomization, a thrombolysis in myocardial infarction (TIMI) risk score of >3, and either an ST-segment deviation of >1 mm or elevated levels of a cardiac biomarker of necrosis.[8] TIMI bleeding is a laboratory-based scale that defines bleeding as major, minor, minimal, or none based on the presence of intracranial hemorrhage, blood loss, hemoglobin concentration, and hematocrit.[9] For patients who had STEMI, patients were enrolled within 12 hours after the onset of symptoms if PCI was planned within 14 days.[8] Exclusion criteria included an increased risk of bleeding, anemia, thrombocytopenia, history of pathological intracranial findings, or the use of a thienopyridine within 5 days before enrollment.[8] Of all the patients enrolled, 99% underwent PCI, 94% received at least 1 stent, and 47% received at least 1 drug-eluting stent.[8]

In total 13,609 patients were enrolled; 10,074 patients had UA/NSTEMI and 3534 patients had STEMI.[8] Patients were randomized and received either prasugrel 60 mg loading dose followed by 10 mg once daily, or clopidogrel 300 mg loading dose followed by 75 mg once daily.[8] The loading dose was given between randomization and 1 hour after leaving the catheterization laboratory, and maintenance therapy was administered for a median duration of 14.5 months.[8] All patients also received aspirin 75 to 162 mg once daily.[8]

The primary study endpoint was a composite of the rate of death from cardiovascular causes, nonfatal myocardial infarction (MI), or nonfatal stroke. The primary endpoint occurred in 12.1% of clopidogrel patients and 9.9% of prasugrel patients (hazard ratio [HR]: 0.81, 95% confidence interval [CI]: 0.73–0.90, $P < 0.001$).[8] A significant reduction in the primary endpoint was seen in both UA/NSTEMI patients ($P = 0.002$) and in STEMI patients alone ($P = 0.02$).[8] Prasugrel therefore showed superior efficacy in reducing the combined rate of cardiovascular death, nonfatal MI, and nonfatal stroke in ACS patients after PCI. This benefit was seen in the prasugrel group at day 3 (5.6% in the clopidogrel group versus 4.7% in the prasugrel group; HR: 0.82, 95% CI: 0.71–0.96, $P = 0.01$) and persisted throughout the study (6.9% in the clopidogrel group versus 5.6% in the prasugrel group; HR: 0.8, 95% CI: 0.7–0.93, $P = 0.003$).[8] Although prasugrel has been shown to reduce the rate of the combined endpoint of cardiovascular-related death, nonfatal MI, and nonfatal stroke, this was largely influenced by the significant reduction in MI in the prasugrel group (9.7% in the clopidogrel group versus 7.4% in the prasugrel group; HR: 0.76, 95% CI: 0.67–0.85, $P < 0.001$).[8] There was no significant difference in the rate of stroke or cardiovascular-related death between the 2 groups[8] (Table I). This suggests that prasugrel may be more beneficial in reducing a patient's risk of MI, but it may not alter one's risk of developing a stroke or other causes of cardiovascular death.

Compared with clopidogrel, prasugrel showed superior efficacy in reducing the combined rate of cardiovascular death, nonfatal MI, and nonfatal stroke in acute coronary syndrome (ACS) patients after percutaneous coronary intervention (PCI).

Table 1. *Efficacy Endpoints of TRITON-TIMI 38 and PRINCIPLE-TIMI 44.*

Trial	Treatment	Endpoint	Prasugrel	Clopidogrel	HR (95% CI)	*P* Value
TRITON-TIMI 38 (n = 13,609)[8]	Prasugrel 60 mg then 10 mg daily or clopidogrel 300 mg then 75 mg daily	CV death, nonfatal MI, nonfatal stroke	9.9%	12.1%	0.81 (0.73–0.90)	<0.001
		CV death	2.1%	2.4%	0.89 (0.7–1.12)	0.31
		Nonfatal MI	7.3%	9.5%	0.76 (0.67–0.85)	<0.001
		Nonfatal stroke	1%	1%	1.02 (0.71–1.45)	0.93
		CV death, nonfatal MI, urgent target-vessel revascularization	10%	12.3%	0.81 (0.73–0.89)	<0.001
		Urgent target-vessel revascularization	2.5%	3.7%	0.66 (0.54–0.81)	<0.001
		Stent thrombosis	1.1%	2.4%	0.48 (0.36–0.64)	<0.001
PRINCIPLE-TIMI 44 (n = 201)[11]	Prasugrel 60 mg then 10 mg daily or clopidogrel 600 mg then 150 mg daily	LD inhibition of platelet aggregation	74.8%	31.8%	–(38–48.4)	<0.0001
		MD inhibition of platelet aggregation	61.3%	46.1%	–(10.6–19.3)	<0.0001

Abbreviations: CI, confidence interval; CV, cardiovascular; HR, hazard ratio; LD, loading dose; MD, maintenance dose; MI, myocardial infarction; PRINCPLE-TIMI, Prasugrel in Comparison to Clopidogrel for Inhibition of Platelet Activation and Aggregation-Thrombolysis in Myocardial Infarction; TRITON-TIMI, Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition with Prasugrel-Thrombolysis in Myocardial Infarction.

> *Prasugrel may be more beneficial in reducing a patient's risk of MI, but it may not alter one's risk of developing a stroke or other causes of cardiovascular death.*

The secondary endpoint of death from cardiovascular causes, nonfatal MI, or urgent target-vessel revascularization was also significantly lower in the prasugrel group (10%) compared with the clopidogrel group (12.3%) (HR: 0.81, 95% CI: 0.73–0.89, $P < 0.001$).[8] Prasugrel was associated with reduced rates of urgent target-vessel revascularization (2.5% prasugrel versus 3.7% clopidogrel; HR: 0.66, 95% CI: 0.54–0.81, $P < 0.001$), and stent thrombosis (1.1% prasugrel versus 2.4% clopidogrel; HR: 0.48, 95% CI: 0.36–0.64, $P < 0.001$).[8] There was also a significant reduction in the number of any target vessel revascularizations (4% versus 10.1%; $P = 0.009$) and coronary artery bypass graft (CABG) surgery (12.5% versus 19.4%; $P = 0.041$) in the prasugrel group.[10]

In subgroup analyses, prasugrel was more beneficial than clopidogrel when used with a glycoprotein IIb/IIIa receptor antagonist (HR: 0.79, 95% CI: 0.69–0.91, $P < 0.001$) or without one (HR: 0.84, 95% CI: 0.72–0.99, $P = 0.03$).[8] Of the 569 patients who underwent PCI without stent placement, the primary composite endpoint of death from cardiovascular causes, nonfatal MI, and nonfatal stroke occurred in 15.6% of nonstent patients compared with 10.8% of stent patients ($P = 0.001$).[10] However, the incidence of the primary endpoint did not differ significantly between the prasugrel (14.2%) and clopidogrel (17.1%) groups (HR: 0.82, 95% CI: 0.53–1.25, $P = 0.27$).[10] Patients were better off with stent placement than without, regardless of the thienopyridine used. Lastly, prasugrel had a greater benefit in patients with diabetes mellitus than in patients without diabetes. Of the 3146 patients with diabetes, 12.2% of prasugrel patients reached the primary endpoint compared with 17% of clopidogrel patients (HR: 0.7, 95% CI: 0.58–0.85, $P < 0.001$).[8,15]

> *Prasugrel had a greater benefit in patients with diabetes mellitus than in patients without diabetes.*

The Prasugrel in Comparison to Clopidogrel for Inhibition of Platelet Activation and Aggregation (PRINCIPLE)-TIMI 44 was a randomized, double-blind, double-dummy, crossover trial that compared prasugrel with clopidogrel in patients undergoing cardiac catheterization for planned PCI.[11] The inclusion criteria were at least 1 lesion amenable to PCI, objective findings of ischemia, or prior PCI or CABG.[11] The exclusion criteria included patients with a planned PCI for immediate treatment of MI, treatment with any thienopyridine within 5 days or a glycoprotein IIb/IIIa inhibitor within 7 days, high bleeding risk, thrombocytopenia, or anemia.[11] Patients received a loading dose of either 60 mg prasugrel or 600 mg clopidogrel about 1 hour before catheterization.[11] Afterward, patients entered the 28-day crossover maintenance phase.[11] Patients received either prasugrel 10 mg daily or clopidogrel 150 mg daily, corresponding to their loading dose assignment, for 14 ± 2 days.[11] Patients then crossed over to the alternate maintenance treatment for another 14 ± 2 days without a washout period.[11] The primary loading dose efficacy endpoint was the inhibition of platelet aggregation with 20 μmol/L of ADP at 6 hours.[11] The primary maintenance dose efficacy endpoint was the inhibition of platelet aggregation with 20 μmol/L ADP after 14 days.[11] Platelet aggregation was measured by light transmission aggregometry and vasodilator-stimulated phosphoprotein.[11] Hyporesponsiveness was defined as inhibition of platelet aggregation <20% with 20 μmol/L ADP.[11]

Approximately 200 patients were randomized with similar maximal platelet aggregation levels of 75.7% in the prasugrel group and 77.0% in the clopidogrel group at baseline ($P = 0.43$).[11] After the loading-dose phase, inhibition of platelet aggregation was significantly larger in the prasugrel group compared with the clopidogrel group (74.8% versus 31.8%; $P < 0.0001$).[11] The maintenance phase also showed greater inhibition of platelet aggregation in the prasugrel group (61.3% versus 46.1%; $P < 0.0001$)[11] (Table 1). Patients in the prasugrel group had more consistent levels of inhibition, less interpatient variability, and less hyporesponsiveness ($P < 0.0001$).[11] Prasugrel patients therefore showed greater platelet aggregation inhibition with less variability and fewer nonresponders.

Compared with clopidogrel, prasugrel patients showed greater platelet aggregation inhibition with less variability and fewer nonresponders.

A lower percentage of nonresponders among prasugrel patients was also shown in a randomized study of 101 aspirin-treated patients with stable coronary artery disease.[12] Patients received one of the following regimens for 28 days: prasugrel 40 mg loading dose followed by 5 mg daily, prasugrel 40 mg loading then 7.5 mg daily, prasugrel 60 mg loading then 10 mg daily, prasugrel 60 mg loading then 15 mg daily, or clopidogrel 300 mg loading then 75 mg daily.[12] Platelet aggregation was assessed using ADP 20 μmol/L within 3 hours of sample collection and was measured by light transmittance.[12] Nonresponse was defined as inhibition of platelet aggregation <20% 4 hours after the loading dose or during maintenance treatment.[12] Both prasugrel loading doses achieved a higher mean inhibition of platelet aggregation than clopidogrel (60.6% and 68.4% versus 30%; $P < 0.0001$), as well as both prasugrel maintenance doses ($P < 0.0001$).[12] There was also a lower percentage of nonresponders after the prasugrel loading dose (3% versus 52%; $P < 0.0001$), and at day 28 there were no nonresponders in the prasugrel group but 45% nonresponders in the clopidogrel group ($P = 0.0007$).[12]

SAFETY

In the Joint Utilization of Medications to Block Platelets Optimally (JUMBO)-TIMI 26 trial, prasugrel was compared with clopidogrel in a phase II, randomized, double-blind, double-dummy safety trial.[13] In total 904 patients undergoing elective or urgent PCI were enrolled for 30 days, and 99% of patients received at least 1 intracoronary stent.[13] All patients received aspirin 325 mg daily and one of the following regimens: clopidogrel 300 mg loading dose followed by 75 mg daily, prasugrel 40 mg loading dose then 7.5 mg daily, prasugrel 60 mg loading dose then 10 mg daily, or prasugrel 60 mg loading dose then 15 mg daily.[13] The primary endpoint was non–CABG-related hemorrhage at 30 days, which was defined as the composite of thrombolysis in myocardial infarction (TIMI) major or minor bleeding.[13] TIMI major bleeding was defined as an intracranial hemorrhage with a drop in hemoglobin of >5 g/dL, and TIMI minor bleeding was defined as a drop in hemoglobin of 3 to 5 g/dL.[13]

There was no statistical difference in non–CABG-related hemorrhage between prasugrel and clopidogrel (1.7% versus 1.2%; HR: 1.42, 95% CI: 0.4–5.08, $P = 0.59$).[13] Major bleeding occurred in 0.5% of prasugrel patients and 0.8% of clopidogrel patients, but this was not a statistically significant difference ($P = 0.544$)[13] (Table 2). The most frequent bleeding site was the vascular access site, and 68% of the bleeding events were due to instrumentation.[13] Furthermore, there was more minimal bleeding in the prasugrel 15 mg group (3.6%) compared with the 10 mg prasugrel group (1.5%), 7.5 mg prasugrel group (2%), and clopidogrel group (2.4%).[13] The 15 mg prasugrel group also had more postdischarge minimal bleeding events (1.2%) compared with the other prasugrel groups (0.5% for each).[13] Although the data suggested that there was a positive association between prasugrel dose and bleeding risk, no statistical analysis was provided to show if the

Table 2. *Safety Endpoints of JUMBO-TIMI 26 and TRITON-TIMI 38.*

Trial	Treatment	Endpoint	Prasugrel	Clopidogrel	HR (95% CI)	P Value
JUMBO-TIMI 26 (n = 904)[13]	Prasugrel 40 mg then 7.5 mg daily or prasugrel 60 mg then 10 mg or 15 mg daily or clopidogrel 300 mg then 75 mg daily	Non-CABG TIMI major or minor bleeding	1.7%	1.2%	1.42 (0.4–5.08)	0.59
		Non-CABG TIMI major bleeding	0.5%	0.8%	0.58 (0.1–3.46)	0.544
TRITON-TIMI 38 (n=13,609)[8]	Prasugrel 60 mg then 10 mg daily or clopidogrel 300 mg then 75 mg daily	Non-CABG TIMI major bleeding	2.4%	1.8%	1.32 (1.02–1.68)	0.03
		Life-threatening bleeding	1.4%	0.9%	1.52 (1.08–2.13)	0.01
		Fatal bleeding	0.4%	0.1%	4.19 (1.58–11.11)	0.002

Abbreviations: CI, confidence interval; HR, hazard ratio; JUMBO-TIMI, Joint Utilization of Medications to Block Platelets Optimally-Thrombolysis in Myocardial Infarction; non-CABG, non–coronary artery bypass graft; TIMI, thrombolysis in myocardial infarction; TRITON-TIMI, Trial to Assess Improvement in Therapeutic Outcomes by Optimizing Platelet Inhibition with Prasugrel-Thrombolysis in Myocardial Infarction.

relationship was significant. According to this study, prasugrel was not associated with a higher bleeding risk compared with clopidogrel.

In the TRITON-TIMI 38 trial, 2.4% of prasugrel patients had at least 1 TIMI major hemorrhage, compared with 1.8% of clopidogrel patients (HR: 1.32, 95% CI: 1.03–1.68, $P = 0.03$).[8] Rates of life-threatening bleeding were greater in the prasugrel group (1.4% prasugrel versus 0.9% clopidogrel; HR: 1.52, 95% CI: 1.08–2.13, $P = 0.01$), as well as the rate of fatal TIMI major bleeding (0.4% prasugrel versus 0.1% clopidogrel; $P = 0.002$)[8] (Table 2). Also, more patients discontinued prasugrel due to hemorrhagic adverse events compared with clopidogrel patients (2.5% prasugrel versus 1.4% clopidogrel; $P < 0.001$).[8]

Rates of life-threatening bleeding were greater in the prasugrel group.

Additional post hoc analysis of the TRITON-TIMI 38 trial showed that certain subgroups of patients did not benefit from the use of prasugrel. Patients with a history of stroke or transient ischemic attack (TIA) had net harm from prasugrel, because 23% of prasugrel patients with a history of stroke or TIA died from any cause compared with 16% of clopidogrel patients (HR: 1.54, 95% CI: 1.02–2.32, $P = 0.04$).[8] In patients with a history of cerebrovascular events, intracranial hemorrhage occurred in 2.3% of prasugrel patients compared with none of the clopidogrel patients ($P = 0.02$).[8] Patients aged >75 years had no net benefit from prasugrel (HR: 0.99, 95% CI: 0.81–1.21, $P = 0.92$), as well as prasugrel patients who weighed >60 kg (HR: 1.03, 95% CI: 0.69–1.53, $P = 0.89$).[8] However, there was no significant difference in the rate of non–CABG-related TIMI major bleeding in those who lacked the above risk factors, regardless of the medication they were assigned to (2% prasugrel versus 1.5% clopidogrel; HR: 1.24, 95% CI: 0.91–1.69, $P = 0.17$).[8]

No net benefit from prasugrel was seen in patients >75 years or in those who weighed <60 kg. Net harm was seen in patients with a history of stroke or transient ischemic attack (TIA).

Other adverse events that were observed in at least 3% of patients in either treatment group were hypertension, hypercholesterolemia, hyperlipidemia, headache, back pain, dyspnea, nausea, dizziness, cough, hypotension, fatigue, and noncardiac chest pain.[4] Severe but rare events included severe thrombocytopenia, anemia, abnormal liver function, allergic reactions, and angioedema.[4]

DISCUSSION

Prasugrel showed greater platelet inhibition with a more rapid onset than clopidogrel. There was a smaller percentage of nonresponders in the prasugrel group, which resulted in less target-vessel revascularization. Overall, prasugrel reduced patients' composite risk of cardiovascular death, nonfatal MI, and nonfatal stroke. The net clinical benefits of prasugrel were also greater in patients with diabetes than without, and in patients who had a nonfatal MI. Of note, the clinical studies reviewed here were short-term trials; therefore, prasugrel's long-term benefits and risks are unknown (Table 3).

The greater effect of prasugrel may be due to its greater inhibition of platelet aggregation; however, factors that affect the conversion of clopidogrel to its active metabolite may also play a role. Conversion of clopidogrel is affected by genetic variants of CYP2C19. In an assessment to determine the association between genetic variants of CYP2C19 and cardiovascular outcomes in clopidogrel patients, levels of active metabolites were determined by liquid chromatography-mass spectrometry, using 20 μM of ADP.[14] Carriers of reduced-function CYP2C19 had 32.4% lower levels of active clopidogrel metabolites ($P < 0.001$), 9% lower absolute reductions in maximal platelet aggregation ($P < 0.001$), and a higher rate of major adverse cardiovascular events than noncarriers (12.1% versus 8%; $P = 0.01$).[14] These results suggest that genetic variants of CYP2C19 play an important role in clopidogrel's ability to inhibit platelet aggregation, and could be a possible reason for clopidogrel's higher percentage of nonresponders. Prasugrel is not affected by pharmacogenetic variants, which may help explain why it has a lower percentage of nonresponders.

There is increasing concern about the effect of proton pump inhibitors (PPI) on clopidogrel metabolism, which may also explain the higher percentage of clopidogrel nonresponders. PPIs inhibit CYP2C19, impeding the conversion of clopidogrel to its active metabolite.[16] Prasugrel, however, is primarily metabolized by CYP3A4 and CYP2B6[4] and is not widely affected by PPI inhibition. In PRINCIPLE-TIMI 44, there were more nonresponders to clopidogrel 6 hours after the loading dose was administered with a PPI compared with those who were not given a PPI (50% versus 18.2%; $P = 0.009$), and there were also more nonresponders after 15 days of maintenance treatment in the PPI group (50% versus 7.9%; $P = 0.012$).[16] In contrast, there was no significant difference in the number of prasugrel nonresponders who were on a PPI versus those who were not 6 hours after the loading dose (0% for both PPI and no PPI); however, there was a significant difference between the 2 groups after 15 days of maintenance treatment (10% versus 0%; $P = 0.025$).[16] In TRITON-TIMI 38, the primary endpoint of cardiovascular death, nonfatal MI, and nonfatal stroke occurred in 11.8% of clopidogrel patients on a PPI versus 12.2% of clopidogrel patients not on a PPI ($P = 0.8$).[16] With prasugrel, 10.2% of patients on a concurrent PPI developed the primary endpoint, compared with 9.7% of patients who were not on a PPI ($P = 0.58$).[16] These results suggested that although PPIs may have increased the number of clopidogrel nonresponders, there was no significant clinical interaction between PPIs, clopidogrel, or prasugrel that would increase patients' risk of cardiovascular death, nonfatal MI, or nonfatal stroke. However, additional data is needed to confirm these results, because other studies have shown an increased risk of adverse events following the use of clopidogrel and PPIs.[16]

Although prasugrel shows greater efficacy with fewer nonresponders, careful patient selection is needed to minimize the risk of major bleeding. Prasugrel has been shown to increase patients' risk of serious, sometimes fatal, bleeding. It should not be used in patients with active pathological bleeding (peptic ulcer or intracranial bleeding),[4] or who have a history of TIA or stroke, since it can result in net harm. Also, prasugrel should not be administered to patients who are likely to undergo urgent CABG surgery, as prasugrel inhibits platelet aggregation for the life of the platelet.[4] Prasugrel should be discontinued at least 7 days before any type of surgery.[4] Prasugrel is not recommended for those >75 years of age or who weigh <60 kg, as these risk factors can increase patients' risk for bleeding. Bleeding is the most commonly reported adverse event and should be suspected in any patient who is hypotensive and underwent a recent surgical procedure, even if there are no signs of bleeding.[4]

Prasugrel should be initiated with a 60 mg oral loading dose, followed by prasugrel 10 mg once daily. Although it is not recommended, practitioners who decide to initiate prasugrel in patients weighing <60 kg should start these patients at 5 mg daily to lower the bleeding risk.[4] The safety and effectiveness of prasugrel 5 mg have not been prospectively studied; however, there may be a positive relationship between prasugrel dose and bleeding risk. Furthermore, although it is not recommended to use prasugrel in patients aged >75 years, it may be beneficial in patients with diabetes or a history of MI, and a lower dose of 5 mg daily may be considered.[4]

Several platelet inhibitors are currently being investigated that could alter the role of prasugrel in the future. Cangrelor, elinogrel, and ticagrelor are all reversible ADP P2Y12 receptor antagonists that allow restoration of platelet function once the medication is discontinued.[17,19,20] Cangrelor is a short-acting, intravenous antagonist that failed to show superior efficacy over clopidogrel in reducing the composite endpoint of death from any cause, MI, or ischemia-driven revascularization in the phase III CHAMPION trial (Cangrelor Versus Standard Therapy to Achieve Optimal Management of Platelet Inhibition).[17] The phase II BRIDGE study (Effectiveness of Bridging Anticoagulation for Surgery) will try to establish cangrelor dosing that will inhibit platelet aggregation by more than 60%.[18] Elinogrel is a competitive antagonist available in both oral and intravenous formulations that is being studied in the phase II trial INNOVATE-PCI (A Randomized, Double-Blind, Active-Controlled Trial to Evaluate Intravenous and Oral PRT060128, a Selective and Reversible P2Y12 Inhibitor, versus Clopidogrel, as a Novel Antiplatelet Therapy in Patients Undergoing Non-Urgent PCI) for safety and efficacy.[19] Lastly, ticagrelor is a direct-acting oral antagonist that significantly reduced the risk of death due to cardiovascular and cerebrovascular causes compared with clopidogrel in the phase III PLATO trial (A Study of Platelet Inhibition and Patient Outcomes).[20] The rate of major, fatal, and life-threatening bleeding did not differ significantly between the 2 groups; however, more ticagrelor patients experienced dyspnea and ventricular pauses.[20] These agents may prove to be safer than prasugrel or clopidogrel due to their reversible antiplatelet mechanisms; however, direct head-to-head studies will be needed to prove their efficacy and safety before determining their proper place in therapy.

CONCLUSION

Prasugrel significantly lowers a patient's risk of major cardiovascular death after PCI, but it also increases the risk of major, sometimes fatal bleeding. Furthermore, there is increased awareness of clopidogrel

Table 3. *Comparative Review of Clopidogrel and Prasugrel.*

	Prasugrel[4]	Clopidogrel[5]
Pharmacokinetics		
Metabolism	1. Rapid hydrolysis through hepatic carboxylesterases	Major pathway: inactive metabolite
	2. 1 CYP450 oxidative step	Minor pathway: active metabolite after 2 CYP450-dependent steps
CYP450 enzymes	3A4, 3A5, 2B6, 2C9, 2C19	3A4, 2C19, 1A2, 2B6
P2Y12 receptor blockade[6]	Complete	Partial
Onset of action	30–60 minutes	2 hours
Contraindications		
Active pathological bleeding	X	X
Prior TIA or stroke	X	
Hypersensitivity		X
Black box warnings		
Bleeding	X	
History of stroke or TIA	X	
Patients aged ≥ 75 years	X	
Need for urgent surgery or CABG	X	
Warnings and precautions		
Bleeding	X	X
CABG-related bleeding	X	
Discontinuation of therapy before surgery	7 days	5 days
Thrombotic thrombocytopenic purpura	X	X
Pharmacogenetics: CYP2C19		X
Special populations		
Hepatic impairment	No dose adjustment in mild to moderate hepatic impairment; severe hepatic impairment has not been studied but may be at higher risk of bleeding	Caution in severe hepatic impairment
Renal impairment	No dose adjustment	Caution in severe renal impairment
Children	Not established	Not established
Pregnancy category	B	B
Breast feeding	Consider risk versus benefit	Discontinue drug

Abbreviations: CABG, coronary artery bypass graft; CYP2C19, cytochrome P450 2C19; CYP450, cytochrome P450; P2Y12, platelet adenosine diphosphate P2Y12 receptor; TIA, transient ischemic attack.

resistance, which may result in patient readmissions due to thrombosis progression. In our institution, approximately 20% of clopidogrel patients are resistant and would benefit from the addition of prasugrel to formulary. Based on our hospital's population and needs, it was recommended that both clopidogrel and prasugrel be made available. Staff education and a protocol will be drafted to ensure proper usage and initiation before PCI, and a 6-month follow-up report will look at all bleeding events to determine if prasugrel use should be continued. Although laboratory testing is available at our institution to detect clopidogrel resistance, it is not the current standard of practice and will not be used as a prescreening tool for prasugrel initiation. Instead, only questionable cases of readmission for thrombosis will be tested to see if treatment failure was due to clopidogrel resistance or due to other causes. If clopidogrel resistance testing becomes the standard of care in the future, the role of prasugrel may change as more data emerges on the rate of clopidogrel resistance at our institution. Due to prasugrel's extensive list of contraindications, precautions, and warnings, prasugrel cannot replace clopidogrel in all patients. Furthermore, there is no data at present regarding the safety or efficacy of switching between clopidogrel and prasugrel once treatment is started. For patients admitted on prasugrel postprocedure, it may be safe and effective to switch patients to clopidogrel, but there is no literature that supports this conversion. Until further studies are done, patients should be continued on their current therapies unless patients are nonresponsive. While prasugrel is associated with a significant bleeding risk, it has been shown to lower the risk of cardiovascular events in specific patient populations, and is associated with fewer nonresponders than clopidogrel. Prasugrel should therefore be considered for formulary addition.

DISCLOSURES

Potential conflict of interest: Nothing to report.

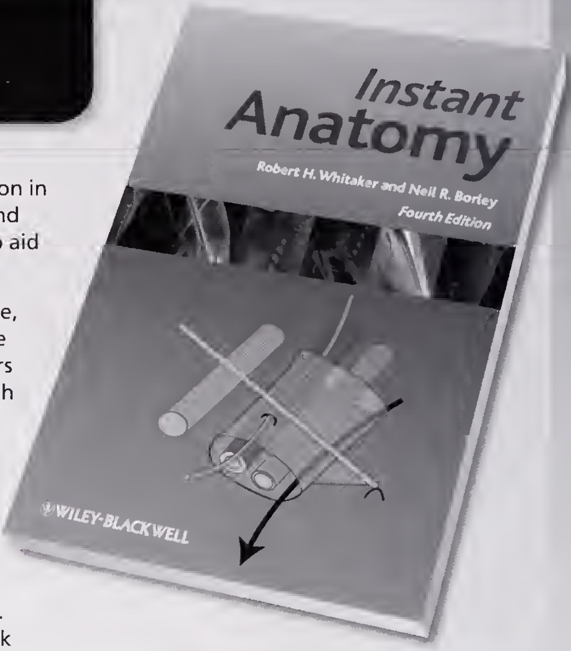
REFERENCES

1. Xu J, Kochanek KD, Tejada-Vera B. Deaths: Preliminary data for 2007. *Natl Vit Stat Rep* 2009; 58: 1–51.
2. Goodman SG, Menon V, Cannon CP, et al. Acute ST-segment elevation myocardial infarction: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). *Chest* 2008; 133(6 suppl): 708S–775S.
3. Harrington RA, Becker RC, Cannon CP, et al. Antithrombotic therapy for non-ST-segment elevation acute coronary syndromes: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). *Chest* 2008; 133(6 suppl): 670S–707S.
4. Effient [package insert]. Indianapolis, IN: Daiichi Sankyo Inc./Eli Lilly and Co.; 2009.
5. Plavix [package insert]. Bridgewater, NJ: Bristol-Myers Squibb/Sanofi Aventis; 2009.
6. Judge HM, Buckland RJ, Sugidachi A, et al. The active metabolite of prasugrel effectively blocks the platelet P2Y12 receptor and inhibits procoagulant and proinflammatory platelet responses. *Platelets* 2008; 19: 125–133.
7. Wallentin L, Varenhorst C, James S, et al. Prasugrel achieves greater and faster P2Y12 receptor-mediated platelet inhibition than clopidogrel due to more efficient generation of its active metabolite in aspirin-treated patients with coronary artery disease. *Eur Heart J* 2008; 29: 21–30.
8. Wiviott SD, Braunwald E, McCabe CH, et al. Prasugrel versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med* 2007; 357: 2001–2015.
9. Rao SV, O'Grady K, Pieper KS, et al. A comparison of the clinical impact of bleeding measured by two different classifications among patients with acute coronary syndromes. *J Am Coll Cardiol* 2006; 14: 809–816.
10. Pride YB, Wiviott SD, Buros JL, et al. Effect of prasugrel versus clopidogrel on outcomes among patients with acute coronary syndrome undergoing percutaneous coronary intervention without stent implantation: a trial to assess improvement in therapeutic outcomes by optimizing platelet inhibition with prasugrel (TRITON)-Thrombolysis in Myocardial Infarction (TIMI) 38 substudy. *Am Heart J* 2009; 158: e21–e26.
11. Wiviott SD, Trenk D, Frelinger AL, et al. Prasugrel compared with high loading- and maintenance-dose clopidogrel in patients with planned percutaneous coronary intervention: the Prasugrel in Comparison to Clopidogrel for Inhibition of Platelet Activation and Aggregation-Thrombolysis in Myocardial Infarction 44 trial. *Circulation* 2007; 116: 2923–2932.
12. Jernberg T, Payne CD, Winters KJ, et al. Prasugrel achieves greater inhibition of platelet aggregation and a lower rate of non-responders compared with clopidogrel in aspirin-treated patients with stable coronary artery disease. *Eur Heart J* 2006; 27: 1166–1173.
13. Wiviott SD, Antman EM, Winters KJ, et al. Randomized comparison of prasugrel (CS-747, LY640315), a novel thienopyridine P2Y12 antagonist, with clopidogrel in percutaneous coronary intervention: results of the Joint Utilization of Medications to Block Platelets Optimally (JUMBO)-TIMI 26 trial. *Circulation* 2005; 111: 3366–3373.
14. Mega JL, Close SL, Wiviott SD, et al. Cytochrome P-450 polymorphisms and response to clopidogrel. *N Engl J Med* 2009; 360: 354–362.
15. Wiviott SD, Braunwald E, Angiolillo DJ, et al. Greater clinical benefit of more intensive oral antiplatelet therapy with prasugrel in patients with diabetes mellitus in the trial to assess improvement in therapeutic outcomes by optimizing platelet inhibition with prasugrel-Thrombolysis in Myocardial Infarction 38. *Circulation* 2008; 118: 1626–1636.
16. O'Donoghue ML, Braunwald E, Antman EM, et al. Pharmacodynamic effect and clinically efficacy of clopidogrel and prasugrel with or without a

proton-pump inhibitor: an analysis of two randomised trials. *Lancet* 2009; 374: 989–997.

17. Harrington RA, Stone GW, McNulty S, et al. Platelet inhibition with cangrelor in patients undergoing PCI. *N Engl J Med* 2009; 361: 2318–2329.
18. Cangrelor [online product information]. Parsippany, NJ: The Medicines Company. http://www.themedicinescompany.com/products_cangrelor. shtml. 2009. Accessed February 2, 2010.
19. Elinogrel [online product information]. South San Francisco, CA: Portola Pharmaceuticals. http://www.portola.com/elinogrel. 2010. Accessed February 12, 2010.
20. Wallentin L, Becker RC, Budaj A, et al. Ticagrelor versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med* 2009; 361: 1045–1057.

DOI:10.1002/MSJ



Instant Anatomy

FOURTH EDITION

Robert H. Whitaker and Neil R. Borley
Selwyn College, University of Cambridge;
Cheltenham General Hospital, UK

Instant Anatomy uses a unique structure to provide core information in an easily accessible 'dip-in' format. It neatly catalogues anatomy and anatomical relationships in a schematic, dictionary-style manner to aid rapid understanding and retrieval of anatomical facts and data.

Now fully updated, this fourth edition of *Instant Anatomy* is precise, concise and easy to navigate, enabling quick look-up and reference points for students, junior doctors and practising clinicians. It covers all the major aspects of anatomy with a consistent, logical approach that pulls together information that would otherwise be spread between sections of a large textbook. Featuring clear illustrations and an expanded surface anatomy section, it presents complex information in a clear and understandable way.

Review of the previous editions

"*Instant Anatomy* is, without doubt, the finest book ever written... It does claim to be unique, and I think it is fair to say that, as a quick revision aid, it is unrivalled. The pictures are simple and ingenious; so much so that one wonders why nobody has written a book like this before."

Medical Student, *Murmur*
(official publication of the University of Cambridge Medical Society).

March 2010 • 9781405199612
242 pages

CONTENTS

Preface to fourth edition
Preface to first edition
Notes on the text

1: Arteries
2: Veins
3: Lymphatics
4: Autonomic nervous system
5: Cranial nerves
6: Peripheral nerves
7: Dermatomes and cutaneous nerve distribution
8: Muscles
9: Joints
10: Ossification times
11: Foramina – skull and spine
12: Position of structures according to vertebral levels
13: Pharyngeal derivatives
14: Surface anatomy and key areas

For more information or to place an order visit: **www.wiley.com**

For illustrations, podcasts, questions and answers, MCQs and PowerPoint presentations on all aspects of anatomy, visit **www.instantanatomy.net**

WILEY-BLACKWELL

MOUNT SINAI JOURNAL OF MEDICINE

VOL. 77, NO. 4
JULY/AUGUST 2010

A Journal of Translational and Personalized Medicine

Issue Theme
Drug Discovery in Academic Settings

Theme Editor
Ravi Iyengar, PhD

Perestroika in Pharma: Evolution of Revolution in Drug Development?

Systems Pharmacology

Discovery of Protein Kinase Targets From Genome-wide mRNA Expression Profiling

Integration of Small Molecule Discovery into Academic Biomedical Research

Pursuit of Personalized Anticancer Therapy by Academia with Biotech/Pharma Industry

Marfan Vascular Disease Drug Therapies: From Mouse Models to Human Patients

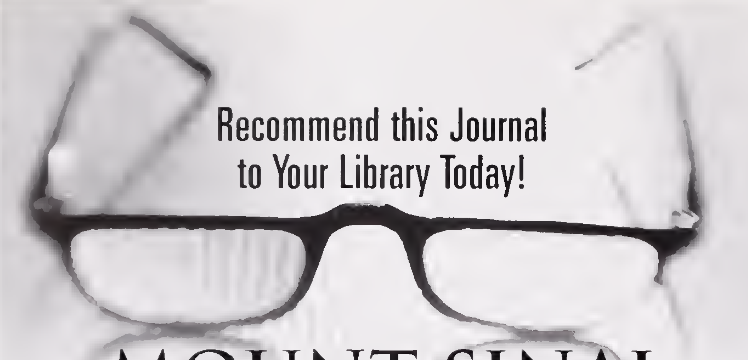
Role of Antibodies in Developing Drugs That Target GPCR Dimers

Pharmacokinetic/ Pharmacodynamic-Driven Drug Development

Special Features

That I Won't Translate: A Family Interpreter in a Multicultural Medical Environment
Ethical Decision-Making in Case of Duchenne Muscular Dystrophy
Mammographic Screening for Breast Cancer

WILEY-BLACKWELL



Recommend this Journal
to Your Library Today!

MOUNT SINAI
JOURNAL OF
MEDICINE

A Journal of Translational and Personalized Medicine

Editor-in-Chief: **Penny A. Asbell, MD, FACS, MBA**
Mount Sinai School of Medicine, New York, NY

Continuously published since 1934 and included in the *Index Medicus* from the very first volume, the **Mount Sinai Journal of Medicine (MSJM)** has a distinguished career as a general medical journal for a worldwide audience.

Dedicated to improving communication among all participants in medicine today, **MSJM** is an excellent source of up-to-date medical information for healthcare clinicians, educators, researchers and policy makers worldwide. **MSJM** will inspire the creative thinking needed to create new ways to understand, prevent, and treat human disease.

Each issue includes several quality, in-depth review articles related to a single theme and several others on subjects of diverse clinical interest. All theme articles are written by renowned authorities in their fields, from both Mount Sinai and elsewhere, under the supervision of Theme Editors.

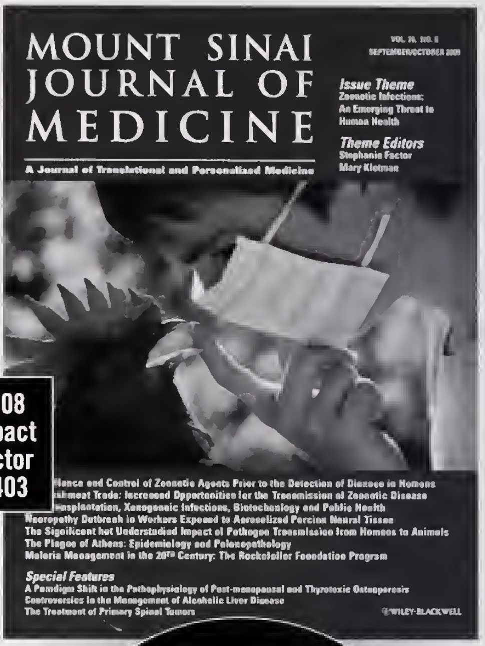
Upcoming Themes for Volume 77:

- Alzheimer's Disease: New Light on Etiology and Management
- Genes, Race, Ethnicity and the Environment
- Endovascular Surgery: Advances in Science and Technique
- Drug Discovery in Academic Settings
- The Obesity Pandemic
- Children's Health and the Environment

ACCESS THESE 2009 EDITORS CHOICE ARTICLES FREE ONLINE

- Advances and Controversies in Assisted Reproductive Technology (76:6)
- Xenotransplantation, Xenogenic Infection, Biotechnology and Public Health (76:5)
- The Student-run Health Clinic: A Novel Arena to Educate Medical Students on Systems-based Practice (76:4)
- Glycemic Targets for Patients with Type 2 Diabetes Mellitus (76:3)
- Neurobehavioral Sequelae of Traumatic Brain Injury (76:2)
- Spectrum of Fibrosing Diffuse Parenchymal Lung Disease (76:1)

Want more? Access a free sample issue online today!



MOUNT SINAI
JOURNAL OF
MEDICINE

A Journal of Translational and Personalized Medicine

Issue Theme
Zoonotic Infections:
An Emerging Threat to
Human Health

Theme Editors
Stephanie Factor
Mary Klotman

Neuropathy Outbreak in Workers Exposed to Aerosolized Porcine Neural Tissue
The Significant but Understudied Impact of Pathogen Transmission from Humans to Animals
The Plague of Athens: Epidemiology and Paleopathology
Malaria Management in the 20TH Century: The Rockefeller Foundation Program

Special Features
A Paradigm Shift in the Pathophysiology of Post-menopausal and Thyrotoxic Osteoporosis
Controversies in the Management of Alcoholic Liver Disease
The Treatment of Primary Spinal Tumors

WILEY-BLACKWELL

2008
Impact
Factor
1.403

Stay on Top
of the Research!

Sign-up for Email Table of Contents Alerts delivered direct to your desktop – FREE!

Simply go to the homepage and click on "Set E-Mail Alert" for automatic table of contents alerts, or create a customized alert based on your own search criteria.

How to Receive Complimentary Online Access for Institutions?

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Library Recommendation Form

To______________________________

Date______________________________

Requested by:________________________

Dept.______________________________

I have reviewed the **Mount Sinai Journal of Medicine** and urge you to add a complimentary 2010 online subscription to our subscription list. Please contact me should you have any questions.

WWW.MOUNTSINAIJOURNAL.ORG

WILEY-BLACKWELL

Enter Your 2010 Subscription Today!

MOUNT SINAI JOURNAL OF MEDICINE

A Journal of Translational and Personalized Medicine

Issue Theme
Type 2 Diabetes Mellitus: New Evidence, More Controversy

Theme Editor
Dina Green

imed at clinicians and educators, the **ount Sinai Journal of Medicine** provides igh-quality, in-depth review articles in very issue, studded with clinical pearls.

our 2010 will include the six 2010 print ssues of the **Mount Sinai Journal of edicine** (Volume 77), plus concurrent text online access to these six issues iley's Internet publishing platform.

dard Personal Subscriptions

0 (U.S./Canada/Mexico)
(Outside North America)

t Sinai Alumni Subscriptions!

(Worldwide)

nded for Institutions in 2010

omplimentary nline Access!

our institution can opt to receive omplimentary online access to the **ount Sinai Journal of Medicine** during 010. When registering for complimentary nline access, your institution can also btain one complimentary print ubscription.

PERSONAL SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 personal subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $105.00 (U.S./Canada/Mexico) ☐ $130.00 (Outside North America)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

Method of Payment (Canada residents please add GST. Personal rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **John Wiley & Sons, Inc.** Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA

Card # ______________________ Exp. Date __________

Signature ______________________

(credit card orders not valid without signature)

Ship To:

Name ______________________

Address ______________________

City ______________ State ________ Zip ________

MSJM0910

MSJM ALUMNI SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 Alumna/Alumnus Subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $40.00 (Worldwide)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

METHOD OF PAYMENT (Alumna/Alumnus rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **The Mount Sinai School of Medicine**. Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA ☐ Other

Card # ______________________ Exp. Date __________

Signature ______________________

(credit card orders not valid without signature)

Ship To:

Name ______________________

Address ______________________

City ______________ State ________ Zip ________

To subscribe online visit: **www.msjmonline.com**

MSJM10

COMPLIMENTARY ONLINE ACCESS FOR INSTITUTIONS

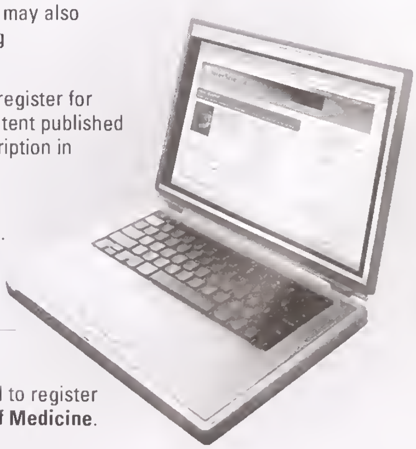
Mount Sinai Journal of Medicine

A Journal of Translational and Personalized Medicine

2010. Volume 77. 6 issues per year. Print ISSN: 0027-2507. Online ISSN: 1931-7581.

Your institution can opt to receive complimentary online access to the **Mount Sinai Journal of Medicine** during 2010. Your institution may also request one complimentary print subscription when registering for complimentary online access.

There is no obligation to subscribe in subsequent years if you register for complimentary access. Institutions retain online access to content published during the complimentary period, with or without a paid subscription in subsequent years.



Dear Librarian,
I have reviewed the scope of Mount Sinai Journal of Medicine.
Please register for complimentary access:

Name: ______________ Department: ______________

Signature: ______________ Date: ______________

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Don't delay, recommend to your library today!

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO. 2277 HOBOKEN, NJ

POSTAGE WILL BE PAID BY ADDRESSEE

JOHN WILEY & SONS INC
ATTN: BETH MOLD, MAIL STOP 8-01
111 RIVER STREET
HOBOKEN, NJ 07030-9441

PLACE
POSTAGE
STAMP
HERE

INES SANTOS
MOUNT SINAI JOURNAL OF MEDICINE
ONE GUSTAV L. LEVY PLACE BOX 1094
NEW YORK, NY 10029-6574

Library Routing Card

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

RECEIVED
AUG 5 2010
Levy Library

Editor-in-Chief

Penny A. Asbell

Deputy Editor

David H. Adams

Associate Editors

Stuart Aaronson
Erwin Bottinger
Dennis S. Charney
Valentin Fuster
Wayne Goodman
Jonathan Halperin
Ravi Iyengar
Paul Klotman
Ihor R. Lemischka
Michael L. Marin
Thomas G. McGinn
Ira Nash
Maria New
Rosamond Rhodes
Hugh Sampson
Mary Sano
Jeffrey H. Silverstein
Albert Siu
Anthony Squire
Fred Suchy
Savio Woo

Editorial Board

William J. Bremner	University of Washington
Thomas M. Coffman	Duke University
Dwight Evans	University of Pennsylvania
Gabriel Haddad	University of California, San Diego
Heidi E. Hamm	Vanderbilt University
Matti Häyry	University of Manchester
Eve J. Higginbotham	Howard University
Katherine A. High	University of Pennsylvania
Tony Hope	University of Oxford
Mary Klotman	Duke University
John H. Krystal	Yale University
Raju Kucherlapati	Harvard Medical School
Robert J. Lefkowitz	Duke University
Robert Lisak	Wayne State University
George Lister	University of Texas
Philip A. Marsden	University of Toronto
Charles Nemeroff	Emory University
Scott Rauch	Harvard Medical School
Bernard Roizman	University of Chicago
Alan F. Schatzberg	Stanford University
Inder Verma	The Salk Institute
Jeffrey A. Whitsett	University of Cincinnati
Allan Wolkoff	Albert Einstein College of Medicine
Noam Zohar	Bar-Ilan University

Theme Editor

Ravi Iyengar, PhD
Chair, Dept. of Pharmacology and Systems Therapeutics
Director, Experimental Therapeutics Institute
Mount Sinai School of Medicine, New York, NY

Editorial Director

Stephen Cohen
stephen.cohen@mssm.edu
212-241-6304

Publisher

John Wiley & Sons
Shawn Morton
Journal Publishing Director, Medicine
smorton@wiley.com

Drug Discovery in Academic Settings

Future *MSJM* Themes and Theme Editors

- Obesity pandemic (Robert Yanagisawa, Derek LeRoith)
- Translational approaches to cancer treatment (William Oh, Steven Burakoff)
- Future of primary care (Thomas McGinn)
- Patient safety (Ira Nash)
- Children's environmental health (Philip Landrigan, Amir Miodovnik)
- Biomarkers in clinical practice (Penny Asbell, Mark Babyatsky)

©2010 **Mount Sinai School of Medicine.** All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, Phone: (978) 750-8400; Fax: (978) 750-4770. Such requests and other permission inquiries should be addressed to the Permissions Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; Phone: (201) 748-6011; Fax: (201) 748-6008; E-mail: permreq@wiley.com. Printed in the United States of America by The Sheridan Press.

Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine (Print ISSN: 0027-2507; Online ISSN: 1931-7581 at Wiley InterScience, www.interscience.wiley.com) is published twelve times per year for the Mount Sinai School of Medicine by Wiley Subscription Services, a Wiley Company, 111 River Street, Hoboken, NJ 07030.

Send subscription inquiries to: Subscription Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-6645.

Offprint sales and inquiries should be addressed to the Reprint Billing Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-8776. All other inquiries should be addressed to the Customer Service Department at (201) 748-6645.

Individual subscription price: $95.00 in North America, $120.00 outside North America. Payment must be made in US dollars drawn on a US bank. Claims for undelivered copies will be accepted only after the following issue has been delivered. For more subscription prices and information, please visit www.mountsinaijournal.org.

Change of address: Please forward six weeks prior to move: present mailing label with change of address. Please enclose a copy of the mailing label or cite your subscriber reference number in order to expedite handling. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits. Send claims in care of: *Mount Sinai Journal of Medicine*, c/o John Wiley & Sons, Inc., Attn: Journals Admin Dept. UK, 111 River Street, Hoboken, NJ 07030.

Cancellations: Subscription cancellations will not be accepted after the first issue has been mailed. Periodicals postage paid at Hoboken, NJ and at additional mailing offices.

Postmaster: Send address changes to *Mount Sinai Journal of Medicine*, Subscription Distribution, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Printed in the United States of America. Single issues from current and recent volumes are available at the current single issue price from cs-journals@wiley.com. Earlier issues may be obtained from Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. Tel: +1 518 537 4700, Fax: +1 518 537 5899, Email: psc@periodicals.com

Advertising inquiries should be addressed to: Advertising Sales Manager, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Phone: (201) 748-6921; Fax: (201) 748-6207. Advertising in this journal is published with a caveat emptor understanding. The Publisher, Editors, and members of the Editorial Board do not imply warranty of products, nor quality, validity, or approval of education material offered by such advertising.

Indexed by: Biological Abstracts® (Thomson ISI), Chemical Abstracts Service/SciFinder (ACS), Current Contents®/Clinical Medicine (Thomson ISI), EMBASE/Excerpta Medica (Elsevier), Index Medicus/MEDLINE/PubMed (NLM), Science Citation Index Expanded™ (Thomson ISI), SCOPUS (Elsevier), Web of Science® (Thomson ISI).

This journal is printed on acid-free paper.

Disclaimer: All articles published, including but not limited to original research, reviews, clinical notes, clinical pearls, prismatic cases, clinician's corner, editorials, reports, and letters, represent the opinions and views of the authors and do not reflect any official policy or medical opinion of the Mount Sinai School of Medicine, the institutions with which the authors are affiliated, or the Publisher unless this is clearly specified. Articles published herein should not be relied upon as recommending or promoting a specific method, diagnosis, or treatment by physicians for any particular patient or disorder. Although the Editors and Publisher believe that drug selections and dosages and the specifications and usage of equipment and devices as set forth herein are in accord with current recommendations and practice at the time of publication, they accept no legal responsibility for any errors or omissions, and make no warranty, express or implied, with respect to material contained herein.

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

CONTENTS
VOL 77, NO 4

Theme Introduction

325 Drug Discovery in Academic Settings
Ravi Iyengar

Theme Articles

327 Perestroika in Pharma: Evolution or Revolution in Drug Development?
Garret A. FitzGerald

333 Systems Pharmacology
Aislyn DW Boran, Ravi Iyengar

345 Protein Kinase Target Discovery from Genome-wide mRNA Expression Profiling
Avi Ma'ayan, John C. He

350 Integration of Small Molecule Discovery into Academic Biomedical Research
Michael Ohlmeyer, Ming-Ming Zhou

358 Pursuit of Personalized Anti-cancer Therapy: Collaboration between Academia and the Biotech/Pharma Industry
Elizabeth Buck, Mark Mulvihill, Kenneth K. Iwata

366 Drug-based Therapies for Marfan Syndrome Vascular Disease: From Mouse Models to Human Patients
Jason R. Cook, Harikiran Nistala, Francesco Ramirez

374 Role of Antibodies in Developing Drugs That Target GPCR Dimers
Chris Hipser, Ittai Bushlin, Achla Gupta, Ivone Gomes, Lakshmi A. Devi

381 Pharmacokinetic/Pharmacodynamic-Driven Drug Development
James M. Gallo

Special Features

389 *That* I Won't Translate: Family Medical Interpreter in Multicultural Environment
Rhona D. Seidelman, Yaacov G. Bachner

394 Ethical Decision-Making in Case of Duchenne Muscular Dystrophy
Leigh Penner, Mollie Cantor, Linda Siegel

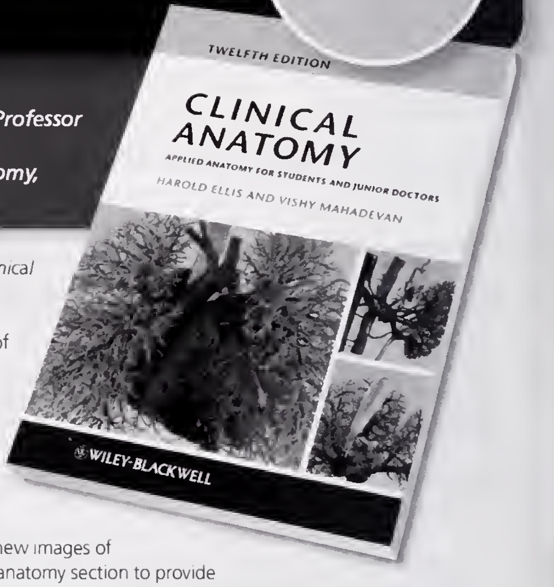
398 Mammographic Screening for Breast Cancer: 2010
Laurie Margolies

Clinical Anatomy

Applied anatomy for students and junior doctors

TWELFTH EDITION

Harold Ellis and Vishy Mahadevan
Guy's, King's and St Thomas' School of Biomedical Sciences; Emeritus Professor of Surgery, Charing Cross and Westminster Medical School, London; Barbers' Company Reader in Anatomy and Professor of Surgical Anatomy, The Royal College of Surgeons of England, London



TWELFTH EDITION
CLINICAL ANATOMY
APPLIED ANATOMY FOR STUDENTS AND JUNIOR DOCTORS
HAROLD ELLIS AND VISHY MAHADEVAN
WILEY-BLACKWELL

Written by one of the great teachers of Anatomy, this 50th anniversary edition of *Clinical Anatomy* continues to provide thousands of medical students, postgraduate trainees and junior doctors across the world with essential anatomical information within a clinical setting. It is particularly appropriate for those preparing for the Membership of the Royal College of Surgeon's Exam.

For this new edition, Professor Harold Ellis is joined, for the first time, by a co-author. Professor Vishy Mahadevan is Barber's Company Reader in Anatomy at the Royal College of Surgeons, a renowned teacher of surgical trainees and is also an examiner for the MRCS Exam. Together they provide a detailed, easy-to-follow structured text suitable for anatomy students and trainees of all levels.

Fully revised and updated, this new edition features new and re-drawn illustrations, new images of normal anatomy (radiographs, CTs and NMRs) to aid understanding, a revised neuroanatomy section to provide greater clarity and an expansion of the lower limb section.

In order to provide revision support, *Clinical Anatomy* has its own website featuring digital flashcards of all the images found in the textbook available to download for free at **www.ellisclinicalanatomy.co.uk**

March 2010 • 9781405186179
470 pages

CONTENTS

Preface to the Twelfth (Jubilee) Edition
Preface to the First Edition
Acknowledgements to the Twelfth (Jubilee) Edition
Acknowledgements to the First Edition

Part 1: The Thorax
Introduction
Surface anatomy and surface markings
The thoracic cage
The lower respiratory tract
The mediastinum
On the examination of a chest radiograph

Part 2: The Abdomen and Pelvis
Surface anatomy and surface markings
The fasciae and muscles of the abdominal wall
Peritoneal cavity
The gastrointestinal tract
The gastrointestinal adnexae: liver, gall-bladder and its ducts, pancreas and spleen
The urinary tract
The male genital organs
The bony and ligamentous pelvis
The muscles of the pelvic floor and perineum
The female genital organs
The posterior abdominal wall
Computed axial tomography

Part 3: The Upper Limb
Surface anatomy and surface markings of the upper limb
The bones and joints of the upper limb
Three important zones of the upper limb: the axilla, the cubital fossa and the carpal tunnel
The arteries of the upper limb
The brachial plexus
The course and distribution of the principal nerves of the upper limb
Compartments of the upper limb
The female breast
The anatomy of upper limb deformities
The spaces of the hand

Part 4: The Lower Limb
The surface anatomy and surface markings of the lower limb
The bones and joints of the lower limb
Three important zones of the lower limb: the femoral triangle, adductor canal and popliteal fossa
The arteries of the lower limb
The veins of the lower limb
The course and distribution of the principal nerves of the lower limb
Compartments of the lower limb

Part 5: The Head and Neck
The surface anatomy of the neck
The thyroid gland
The palate
The tongue and floor of the mouth
The pharynx
The larynx
The salivary glands
The major arteries of the head and neck
The veins of the head and neck
The lymph nodes of the neck
The cervical sympathetic trunk
The branchial system and its derivatives
The surface anatomy and surface markings of the head
The scalp
The skull
The paranasal sinuses (accessory nasal sinuses)
The mandible
The vertebral column

Part 6: The Nervous System
The brain
The spinal cord
The cranial nerves
The special senses
The autonomic nervous system

Glossary of eponyms
Index

For more information or to place an order visit:
www.ellisclinicalanatomy.co.uk

WILEY-BLACKWELL

COMING SOON!
Download the FREE digital flashcards from *Clinical Anatomy*, at **www.ellisclinicalanatomy.co.uk**

THEME INTRODUCTION

Drug Discovery in Academic Settings

Ravi Iyengar, PhD

Department of Pharmacology and Systems Therapeutics, Experimental Therapeutics Institute,
Mount Sinai School of Medicine, New York, NY

Over the past 50 years, the process of drug discovery and development has largely been conducted in an industrial setting, both in large multinational pharmaceutical companies and in small biotechnology companies that focus on the early phases of drug discovery. Overall, drug development has been highly successful, with life-threatening diseases such as hypertension and infectious diseases such as HIV converted to chronic manageable conditions. Diseases such as ulcers that once had to be treated surgically are now most commonly treated with drugs. Despite these successes, there is a common recognition that drug discovery and development has not kept up with the growth of knowledge in biomedical sciences, and the rate of approval of new drugs by the US Food and Drug Administration (FDA) has slowed to a trickle. In part, the problems in drug discovery and development arise from the fact that current drug development is focused on complex diseases such as cancers, psychiatric disorders, and type 2 diabetes. These diseases are more appropriately thought of as a set of diseases with varying molecular origins that result in manifestation of similar clinical characteristics. This complexity requires that basic research on the origins and progression of diseases be tightly integrated with the drug discovery and development process. One natural habitat for such integration is an academic research unit. However, due to the explosive growth in studies at the molecular and cellular level, most academic departments of pharmacology have focused on basic mechanistic studies and there has been little academic research on the process of drug discovery and development for complex diseases. This situation is starting to change: The National Institutes of Health has funded several centers for screening of new molecules that can then be developed as drugs, and many academic medical centers and universities have started interdisciplinary programs in drug discovery and development. This is an encouraging trend, although it is too early to tell whether such efforts will have a systemic effect and speed up the process of drug discovery and development.

Analyzing and understanding the various steps in the drug discovery and development process within the context of academic institutions is likely to be a useful exercise, because it will allow us to identify loci where synergies can occur. Such identification can then serve as a criterion for resource allocation that promotes interdisciplinary research, which in turn can lead to synergistic processes in drug discovery and development. In this set of reviews, we have focused on the early phase of drug discovery and development in an academic context. The first article, a commentary by Fitzgerald, analyzes the various issues we need to deal with if the focus of drug discovery turns from industry to academia. The second article, by Boran and Iyengar, describes a conceptual framework of how systems-biology approaches can be used to tackle the problem of drug discovery for complex diseases. In the third article, by Ma'ayan and He, protein kinases are discussed as an example to describe how computational systems-biology approaches can help us analyze large datasets to identify and rank drug targets. The next article, by Ohlmeyer and Zhou, focuses on high-throughput screening of small-molecule libraries against drug targets and the steps involved in taking a molecule identified in such a screen to the level of a lead compound. In the fifth article, Buck, Mulvihill, and Iwata describe how cooperation between academia and biotechnology companies has catalyzed the development of targeted cancer drugs. This article provides a useful picture

Address Correspondence to:

Ravi Iyengar
One Gustave L. Levy Place
New York, NY
Email: ravi.iyengar@mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20190

© 2010 Mount Sinai School of Medicine

of how the drug discovery and development process has moved from academia to industry in the past. It is anticipated that this process will have to be reengineered to enhance the rate of drug discovery for complex processes.

The next 2 articles describe new basic research approaches for drug discovery. Cook, Nistala, and Ramirez use Marfan syndrome as an example to describe how drug therapy can be used to treat congenital disorders and how the drug therapy for Marfan syndrome, a genetic disease, has opened up avenues to treat acquired connective-tissue disorders. Devi and colleagues describe a potentially fascinating approach for development of a new class of drugs for a family of well-established drug targets: the G-protein-coupled receptors. They propose that antibodies that selectively recognize receptor dimers could function as drugs. Although antibodies against receptor tyrosine kinases are in clinical use as cancer drugs, no antibody drugs have been characterized for G-protein-coupled receptors. Drugs with specificity against types of receptor homodimers and heterodimers, if developed, could be highly selective for complex diseases.

The last article in the series deals with the critical role of pharmacokinetic and pharmacodynamic studies in drug development. As Gallo rightly points out, up to now pharmacokinetic studies have largely focused on obtaining data that is required for regulatory approval by the FDA rather than on the relationships between organismal-level pharmacokinetic characteristics, drug distribution in the tissue of relevance, and mechanisms of drug action. These types of integrated studies could lead to early estimations of efficacy that can accelerate the drug-development process. Indeed, a tight iterative integration between phase I clinical trials and pharmacodynamic studies could be a locus where the emergent synergy speeds up the drug-development process. Iterative integration of clinical trials with the early phases of drug development is a process that needs further analyses leading to a de novo design of drug-development process flow. I hope this set of articles will provide an impetus for such translational analyses.

DOI:10.1002/MSJ

Perestroika in Pharma: Evolution or Revolution in Drug Development?

Garret A. FitzGerald

University of Pennsylvania School of Medicine, Philadelphia, PA

OUTLINE

MIND THE GAP: BRIDGING THE DEFICIT IN HUMAN CAPITAL
EXPAND THE PRECOMPETITIVE SPACE: PUBLIC-PRIVATE PARTNERSHIPS AND REFORM OF INTELLECTUAL PROPERTY
LAY THE FOUNDATIONS OF A GLOBAL MODULAR NETWORK: A COMPLIANT INFRASTRUCTURE FOR DATA SHARING
CAN VENTURE BE ADVENTUROUS? NEW MODELS FOR FINANCING DRUG DEVELOPMENT
CONCLUSION

ABSTRACT

New-drug approvals have remained roughly constant since 1950, while the cost of drug development has soared. It seems likely that a more modular approach to drug discovery and development will evolve, deriving some features from the not-for-profit sector. For this to occur, we must address the deficit in human capital with expertise in both translational medicine and therapeutics and also in regulatory science; utilize regulatory reform to incentivize innovation and the expansion of the precompetitive space; and develop an informatics infrastructure that permits the global, secure, and compliant sharing of heterogeneous data across academic and industry sectors. These developments, likely prompted by the perception of crisis rather than opportunity, will require linked initiatives among academia, the pharmaceutical industry, the US National Institutes of Health, and the US Food and Drug Administration, along with a more adventurous role for venture capital. A failure to respond threatens the United States' lead in biomedical science and in the development and regulation of novel therapeutics. ***Mt Sinai J Med 77:327–332, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: drug discovery, human capital, regulatory science, therapeutics, translational medicine.

Address Correspondence to:

Garret A. FitzGerald
University of Pennsylvania
School of Medicine
Philadelphia, PA
Email: garret@upenn.edu

The declining rate of new-drug approvals suggests that the current approach to drug development is unsustainable.[1] Although a revolution in drug discovery has yielded an abundance of rationally selected drug targets, drug development remains unreformed. This has coincided with a dramatic decline in the growth of prescription-drug sales in the United States (from >15% in 2000 to ~1% in 2007) and a shift toward generic products that now account for ~60% of the $290 billion spent annually in the United States on prescription drugs.[2] In the decade from 1996 to 2006, prescription-drug costs rose from comprising 10% to comprising 18% of healthcare costs.

Among pharmaceutical companies, a strategy of merger–increasing 10-fold in the past 5 years to a deal value of $220 billion in the first quarter of 2009–has delivered some short-term financial gains (almost 90,000 jobs shed in 2009). However, mergers have failed to foster innovation, sustain creativity, and accelerate the delivery of new drugs.

The pharmaceutical industry (Pharma) has shifted its focus toward cancer, driven in part by the perception of an uneven regulatory landscape, at the expense of what remains the leading cause of death: cardiovascular disease.[3] This shift has also affected other areas, such as neurological and psychiatric diseases. However, reimbursement for expensive biopharmaceuticals that confer minor incremental life extension in cancer is now questioned.

The success of the National Institute for Health and Clinical Excellence (NICE) in promoting cost containment in the United Kingdom suggests that a similar approach will ultimately integrate with initiatives in comparative effectiveness further to depress US prescription-drug pricing. An interesting aspect of NICE's activities has been to render reimbursement conditional on demonstrable efficacy;

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20198

© 2010 Mount Sinai School of Medicine

in other words, it serves to foster innovation. The passage of healthcare reform is likely to accelerate these changes in the United States.

Combination of drugs with genomic analysis and quantitative biomarkers of drug response as medicine becomes "personalized" might theoretically defend price erosion of branded drugs and add value to their generic equivalents; however, it is arguable how effectively this will attract incremental reimbursement.[4]

These developments have prompted cogitation and analysis by groups such as the Institute of Medicine Forum on Drug Discovery, Development, and Translation (http://www.iom.edu/CMS/3740/24155.aspx); an increasing number of mostly small (<\$10 million per year) and conventional deals to foster industry-academia interaction; limited access by academics to drugs that have previously failed in development by Pharma; and attempts to pursue common interests, such as biomarkers, by consortia of companies, academics, and the US Food and Drug Administration (FDA).

Can the present business model adapt and survive with a little perestroika and a touch of glasnost? Alternatively, despite power, money, and global reach, might the vertically integrated structure of Pharma–like General Motors, Lehman Brothers, and the Soviet Union–ultimately collapse quite abruptly under the weight of its own contradictions?

How might industry, academia, and the FDA position themselves to anticipate such transition to a more global, less-integrated, and modular approach to drug discovery and development?

MIND THE GAP: BRIDGING THE DEFICIT IN HUMAN CAPITAL

The number of people who can project basic science across the translational divide into mechanistic studies of drug action in humans has declined alarmingly in industry, academia, and the FDA.

> *The number of people who can project basic science across the translational divide into mechanistic studies of drug action in humans has declined alarmingly in industry, academia, and the FDA.*

This reflects both the demise of education in the systemic disciplines of pharmacology and physiology relative to reductionist molecular approaches and the failure of clinical pharmacology to sustain traction in academia. Although a team approach to assembly of the skill sets necessary to guide a drug from preclinical proof of concept (POC) to rational and stratified selection of dose in humans is possible, those few individuals who integrate such skills within their own experience can prove catalytic to the process.

It is the responsibility of academia to provide such interdisciplinary training; however, the experience should include exposure to drug development in a commercial setting and to the demands and process of regulatory science. Some initiatives have been launched to address this gap in manpower. The Wellcome Trust has funded 4 centers of excellence in translational medicine and therapeutics focused on training; each academic center has an industrial partner that contributes expertise to the teaching program and an environment where trainees can be exposed to drug development in the setting of a for-profit company.[5] More recently, the Medical Research Council in the United Kingdom has funded programs designed to address a skills gap in both clinical pharmacology and pathology in that country.[6] In the United States, the consortium of the National Institutes of Health (NIH) Clinical and Translational Science Awardees (CTSA) has designated education as the second of its strategic goals and so-called "T1" research–that which spans the translational divide between preclinical models and deep phenotyping in relatively small numbers of people–as its fifth strategic goal.[7] The consortium has also launched initiatives to foster interaction among the academic sector, industry, and the FDA focused on improving the efficiency of the development of novel therapeutics.[8]

If such arduous training is to attract bright graduates, then the discipline must be branded and promoted. This requires use of a common nomenclature by the stakeholders, the funding of training programs, and the evolution of career structures. Old brands–clinical pharmacology and experimental medicine–might be difficult to resuscitate and fail to capture, respectively, the breadth of training or the focus on therapeutics now required.[9] New ones, like CTS embrace this activity, but within a broad church where coexisting constituents–basic scientists, clinical epidemiologists, clinical trialists, and health-service researchers–crowd out the pursuit with no name.

It is periodically fashionable for Pharma to establish clusters of physician-scientists charged with exploring mechanisms to attenuate the risk of late-stage failure and to speed the efficiency of drug development. However, they are poorly positioned to argue against the pressures exerted by marketing or venture capital to accelerate from initial evidence of human drug tolerability into large-scale trials.

If translational medicine is to integrate with knowledge of therapeutics and flourish in academia, it must be fostered as an academic discipline with revision of promotional and funding timelines to accommodate the nature of translational research. The objective here is not to turn universities into drug companies; rather, it is to increase the number of translational scientists who can pursue research between POC in cells and model systems and the study of mechanism–physiology, disease, and drug action–in humans.[10] Some of these individuals will remain in academia, but there is a need for such interdisciplinary expertise in Pharma, biotech, venture industries, and the FDA.

EXPAND THE PRECOMPETITIVE SPACE: PUBLIC-PRIVATE PARTNERSHIPS AND REFORM OF INTELLECTUAL PROPERTY

Both industry and academia have outmoded and unrealistic approaches to intellectual property (IP). This has constrained interaction between the 2 sectors at an unmeasured opportunity cost.

> *Both industry and academia have outmoded and unrealistic approaches to intellectual property.*

Ironically, the crisis in Pharma has prompted short-term, rather than strategic, thinking. Thus, strategically beneficial investments, such as in precompetitive collaborations, are constrained. Obvious areas in which to expand IP free space are systems pharmacology and physiology[1,11] and the elucidation of off-target effects of drug action, both ripe for funding via public-private partnerships (PPPs). Another example is the sharing of information across companies and sectors to seek genetic predisposition to adverse effects.[12] For these initiatives to develop and expand, the FDA might create a "safe haven" for systems physiology and pharmacology, akin to its approach to pharmacogenomics. Presently, companies are disincentivized to pursue or sanction preapproval evaluations of investigational new drugs unrelated to their primary indication.

> *Presently, companies are disincentivized to pursue or sanction preapproval evaluations of investigational new drugs unrelated to their primary indication.*

Increased interaction between academia and the FDA will also revise the balance between the proprietary interests of the sponsor and the public. If incented to do so by the provision of funds, academia can augment the expertise of the FDA in regulatory science by providing critical mass in emerging sciences, a place to pursue aspects of drug action that arise during the approval process, and a site for bidirectional educational interchange. Such an initiative was suggested in the report of a subcommittee of the FDA Science Board.[13] It would supplant the Critical Path initiative that efficiently identified gaps but was never sufficiently resourced to engage academia in a comprehensive manner. An initiative such as that envisaged by the Science Board–a "Jet Propulsion Lab" for the FDA–would add value to regulatory science by drawing on 2 other endeavors resourced from the public purse: the academic sector, from which the FDA would draw expertise, and the NIH, which could help by deploying its well-established peer-review system to assist the FDA in the assessment of proposals and the allocation of funds.

As the need for laws governing IP are fundamental to the development of profitable drugs, their reform and its restriction to domains of established value seems timely. Expansion of the precompetitive space was among the topics of consideration at meetings in the NIH and the Wellcome Trust in the first half of 2010.

LAY THE FOUNDATIONS OF A GLOBAL MODULAR NETWORK: A COMPLIANT INFRASTRUCTURE FOR DATA SHARING

Harnessing and integrating the expertise of the best groups across the world at the various stages of drug discovery and development will rely on a secure, compliant, globally accessible informatics infrastructure that is sufficiently flexible to permit access by industry, academic, and regulatory partners. Large Pharma already communicates such information intramurally to some degree, whereas academic investment has lagged and been uncoordinated, even within institutions. Upgrading information systems has been a priority of the CTSA initiative, but in and of itself this funding system is far from sufficient to address the problem. Here again, industry and academia might combine effectively. Perhaps one model for such an interaction is the Translational Medicine Research Collaboration involving Pfizer, 4 UK National Health Service trusts, Scottish Enterprise, and 4 Scottish universities.[14]

Meantime, companies might consider rendering aspects of their information systems as open source to establish and disseminate best-practice platforms.[15] Development of such an infrastructure and the attendant international standards for annotation, compliance, and security might also be financed by PPPs.

The relevance of the future development of therapeutics to the public health and wealth should be sufficient incentive for governments to engage.

CAN VENTURE BE ADVENTUROUS? NEW MODELS FOR FINANCING DRUG DEVELOPMENT

Here, some version of the future may beckon from the not-for-profit sector.[1] Although altruism and the diminished perception of likely profit are relevant, the creation of a sufficient fund has engaged leading academic groups to interact with industrial partners, relax or postpone their claims of IP, hand off data sets in a timely fashion, and share academic credit. Impressive examples include Medicines for Malaria Venture,[16] the TB Alliance,[17] and the Institute for OneWorld Health.[18]

Another model is to use prizes or funds to foster innovation. Such an approach has been used quite broadly across industries to foster innovation and discovery.[19] Such incentives have been launched in many countries[20,21] and in targeted areas, such as genomics,[22] to hasten the personalization of medicine. The Galen Prize (Prix Galien) is designed to reward innovation within Pharma and the biotech and medical-device industries.[23]

The academic sector particularly needs to be empowered if it is to contribute more substantially to drug and device development. Reform of funding cycles and promotional expectations to facilitate translational science have been noted. Additionally, academia generally lacks the capacity to perform the medicinal chemistry to harvest the potential therapeutic benefit of its biological discoveries. Few US academic departments of chemistry are incented to fulfill this goal. Here again, there is opportunity to interact with Pharma or with academic chemistry departments in countries where cost structures are lower and expectations of chemistry departments are different from those in the United States and other developed countries.

This issue also speaks to the lack of funding available within academia to pursue the development of a therapeutic. NIH funding structures are not suited to support the necessary replication of POC studies, pharmacokinetics studies and modeling, the extramural toxicology screens, and the requisite medicinal chemistry and scaling required of such an endeavor. There exist scattered, useful, and underexploited initiatives, such as the NIH Rapid Access to Interventional Development (RAID) program.[24] However, these are relatively inflexible (fixed funding cycles), small-scale initiatives that do not address the increasing tendency of venture capital to postpone investment until new therapeutic entities have reached the POC stage in humans.

Given the objectives of the CTSA program and the prioritization of both translational and personalized medicine by the new director of the NIH,[25] perhaps now is the time to develop an integrated initiative to empower the emerging potential of the academic sector to contribute to drug and device development. Meantime, an interesting initiative in the private sector is the codevelopment with academic investigators of an early lead with the first option of buyback at a later, more advanced stage of development.[26]

The commonly cited figure of $1 billion to $4 billion to develop a drug[27] is a disincentive to those contemplating seminal reform. However, this number is only partially relevant to consideration of a substantial revision of the business model, as it builds in the price of failure pursuant of the model to be replaced. Experience from the not-for-profit sector suggests that the cost of drug development in a devolved model, integrating modules across continents and sectors, might be considerably lower if accompanied by a deferred model of reward based on the distribution of risk. Even within the parameters of the existing model, the cost and likelihood of bringing drugs to market successfully vary considerably between large and small molecules and their intended therapeutic indications.[28]

CONCLUSION

The rate at which new molecular entities have been approved has remained essentially constant since

1950, a period during which the cost of discovering and developing such therapeutics has soared[29]; this is clearly unsustainable. And it is all the more so given the political pressures to reduce the pricing of patented drugs. These pressures are likely to hasten the disintegration of the vertically integrated large Pharma companies and prompt a move to a more modular approach with shifting engagement of expertise across sectors and geography, as is currently exemplified in the altruistic sector.

To project this model into the for-profit sector will require a substantial revision of the ground rules regarding IP and imaginative and transparent approaches to the disclosure and possibly the "management" of conflict of interest. The FDA might usefully incentivize true innovation by the use of accelerated approval strategies, harvest the value of academia with collaborative centers of excellence, and foster an earlier, more thorough understanding of drug action by creating a "safe haven" for systems physiology and pharmacology, spanning the translational divide. Finally, the NIH might aggregate and expand its initiatives relevant to drug development and collaborate with academia to promote training and career development in translational medicine and therapeutics.[29,30]

Although drug development is presently undergoing a process of modest and gradual reform, rapidly evolving economic and political pressures[31] may converge to prompt a much more disruptive initiative. This, to the great surprise of many executives at the time, previously occurred in the publishing, airline, music, and financial-services industries.

> "Go on failing. Go on. Only next time, try to fail better."
>
> Samuel Beckett

ACKNOWLEDGMENTS

The author is grateful for the helpful comments of Tilo Grosser, John Hogenesch, and Kevin Horgan.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. FitzGerald GA. Drugs, industry and academia. *Science* 2008; 320: 1563.
2. Pharmaceutical ills. "Lex" column. *Financial Times*. Published April 1, 2008. http://www.ft.com/cms/s/3/f0365ebc-ffc6-11dc-825a-000077b07658.html. Accessed June 15, 2010.
3. Taroncher-Oldenburg G. The heart of translation. *SciBX* 2009;doi:10.1038/scibx.2009.167. http://www.nature.com/scibx/journal/v2/n5/full/scibx.2009.167.html. Published February 5, 2009.
4. Davis JC, Furstenthal L, Desai AA, et al. The microeconomics of personalized medicine: today's challenge and tomorrow's promise. *Nat Rev Drug Discov* 2009; 8: 279–286.
5. Wellcome Trust. Translational Medicine and Therapeutics programmes. http://www.wellcome.ac.uk/Funding/Biomedical-science/Grants/PhD-programmes-and-studentships/WTD027975.htm.
6. United Kingdom Medical Research Council. Clinical Pharmacology and Pathology Fellowship Programmes. http://www.mrc.ac.uk/Fundingopportunities/Fellowships/Clinicalpharmacologypathology/index.htm.
7. CTSA Consortium. Clinical and Translational Science Awards program. http:/www.ctsaweb.org/index.cfm?fuseaction=quicklink.showSGC.
8. National Institutes of Health National Center for Research Resources. CTSA Industry Forum. http://guest.cvent.com/EVENTS/Info/Summary.aspx?e=0b0d14f7-a92c-4e39-8096-82fc07cad7c5.
9. FitzGerald GA. Clinical pharmacology or translational medicine and therapeutics: reinvent or rebrand and expand? *Clin Pharmacol Ther* 2007; 81: 19–20.
10. FitzGerald GA. Opinion: anticipating change in drug development: the emerging era of translational medicine and therapeutics. *Nat Rev Drug Discov* 2005; 4: 815–818.
11. Schadt EE, Friend SH, Shaywitz DA. A network view of disease and compound screening. *Nat Rev Drug Discov* 2009; 8: 286–295.
12. International Serious Adverse Event Consortium. http://www.saeconsortium.org.
13. US Food and Drug Administration. FDA science and mission at risk: report of the FDA Science Board's Subcommittee on Science and Technology. http://energycommerce.house.gov/Press_110/022508.ScienceBoardReport.EstimatedResources.pdf. Published February 25, 2008.
14. Translational Medicine Research Collaboration. http://www.tmrc.co.uk.
15. Barnes MR, Harland L, Foord SM, et al. Lowering industry firewalls: pre-competitive informatics initiatives in drug discovery. *Nat Rev Drug Discov* 2009; 8: 701–708.
16. Medicines for Malaria Venture. http://www.mmv.org. Accessed June 15, 2010.
17. TB Alliance. http://www.tballiance.org/home/home.php.
18. Institute for OneWorld Health. http://www.oneworldhealth.org.
19. Innovation Prize Central. http://innovationprizecentral.com.
20. The Science|Business Network. Nominations open for £20,000 bioscience innovation prize. http://bulletin.sciencebusiness.net/sb/login.php?page=/ebulletins/showissue.php3?page=/548/art/15322. Published October 8, 2009.
21. European Patent Office. European Inventor Award. http://www.epo.org/topics/innovation-and-economy/european-inventor.html.

22. The X Prize Foundation. Archon X Prize for Genomics. http://genomics.xprize.org.
23. Prix Galien USA. The Galen Prize. http://submission.prix-galien-usa.com.
24. US National Institutes of Health Roadmap. NIH Rapid Access to Interventional Development (NIH-RAID). http://nihroadmap.nih.gov/raid.
25. Kaiser J. Collins sets five themes for NIH. http://blogs.sciencemag.org/scienceinsider/2009/08/collins-sets-fi.html. Published August 17, 2009.
26. Symphony Capital, LLC. http://www.symphonycapital.com.
27. Vernon JA, Golec JH, Dimasi JA. Drug development costs when financial risk is measured using the Fama-French three-factor model. *Health Econ*;doi:10.1002/hec.1538.
28. DiMasi JA, Feldman L, Seckler A, Wilson A. Trends in risks associated with new drug development: success rates for investigational drugs. *Clin Pharmacol Ther* 2010; 87: 272–277.
29. Munos B. Lessons from 60 years of pharmaceutical innovation. *Nat Rev Drug Discov* 2009; 8: 959–968.
30. Skarke C, FitzGerald GA. Training translators for smart drug discovery. *Sci Transl Med* 2010; 2: 26 cm12.
31. Future budget deficits almost entirely due to rising private sector health care costs [press release]. Washington, DC: Center for Economic and Policy Research; August 12, 2009. http://www.cepr.net/index.php/press-releases/press-releases/future-budget-deficits-almost-entirely-due-to-rising-private-sector-health-care-costs.

MOUNT SINAI JOURNAL OF MEDICINE 77:333–344, 2010

Systems Pharmacology

Aislyn D. W. Boran, PhD, and Ravi Iyengar, PhD

Mount Sinai School of Medicine, New York, NY

OUTLINE

ANALYZING MECHANISMS OF DISEASE: THROUGH REGULATORY NETWORKS
ANALYZING DRUG-TARGET RELATIONSHIPS WITH REGULATORY NETWORKS
TARGETED THERAPIES: ADVANTAGES AND PITFALLS
COMPUTATIONAL METHODOLOGIES IN DRUG DISCOVERY
CONCLUSION
ACKNOWLEDGMENTS

ABSTRACT

We examine how physiology and pathophysiology are studied from a systems perspective, using high-throughput experiments and computational analysis of regulatory networks. We describe the integration of these analyses with pharmacology, which leads to new understanding of drug action and enables drug discovery for complex diseases. Network studies of drug-target relationships can serve as an indication on the general trends in the approved drugs and the drug-discovery progress. There is a growing number of targeted therapies approved and in the pipeline, which meets a new set of problems with efficacy and adverse effects. The pitfalls of these mechanistically based drugs are described, along with how a systems view of drug action is increasingly important to uncover intricate signaling mechanisms that play an important part in drug action, resistance mechanisms, and off-target effects. Computational methodologies enable the classification of drugs according to their structures and to which proteins they bind. Recent studies have combined the structural analyses with analysis of regulatory networks to make predictions about the therapeutic effects of drugs for complex diseases and possible off-target effects. ***Mt Sinai J Med 77:333–344, 2010.*** **©** *2010 Mount Sinai School of Medicine*

Key Words: drugome, signaling networks, systems biology, systems pharmacology, targeted therapy.

Classical studies such as the development of receptor theory by Clark[1] and Black,[2,3] followed by analyses that distinguished between competitive and noncompetitive inhibition, served as the initial insight on the molecular mechanisms of drug action.[4] The influence and relevance of receptor theory in modern pharmacology are due to the large number of drugs that target membrane receptors, mostly G-protein-coupled receptors (GPCRs). A majority of the remaining drugs are enzyme inhibitors, which are generally substrate-based analogs, transition-state analogs, or allosteric inhibitors, designed based on the structures of the substrate and the substrate-binding pocket to bind reversibly or irreversibly. Over the past decade, genomics, proteomics, network analyses, and other high-throughput studies have led to a wealth of "systems-level" information. Through combinations of the human genome studies, high-throughput technologies, and structural and biochemical studies, the number of well-characterized, druggable targets is constantly increasing.[5,6] Hence, the drug pipeline has grown, with a focus on complex, multigenic diseases resulting in the appearance of targeted therapies and biological therapeutics, such as kinase inhibitors and monoclonal antibody therapies. Despite these advances, the overall rate of development of new drugs has greatly decreased.

There are more than 1500 drugs approved by the US Food and Drug Administration (FDA), but only approximately 400 unique proteins are targeted by these approved drugs (the "drugome") (Figure 1). The diseaseome is defined as the current ~3700 unique entries in the Online Mendelian Inheritance

Address Correspondence to:

Ravi Iyengar
Department of Pharmacology
and Systems Therapeutics
Mount Sinai School of Medicine
New York, NY
Email: ravi.iyengar@mssm.edu

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20191

© 2010 Mount Sinai School of Medicine

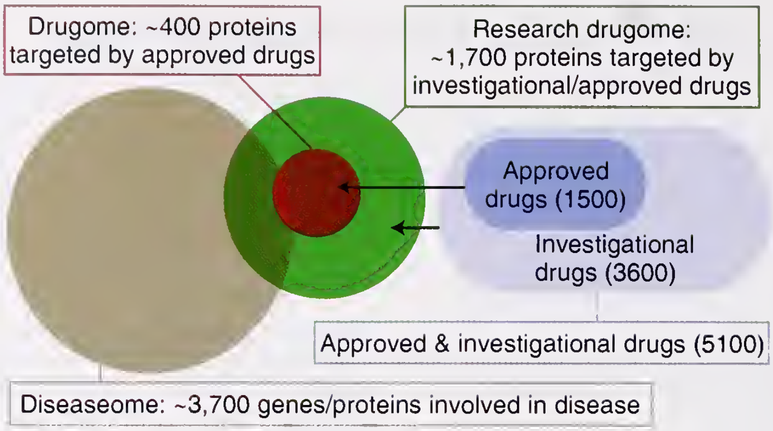

Drugome: ~400 proteins targeted by approved drugs
Research drugome: ~1,700 proteins targeted by investigational/approved drugs
Approved drugs (1500)
Investigational drugs (3600)
Approved & investigational drugs (5100)
Diseaseome: ~3,700 genes/proteins involved in disease

Fig 1. Relationships between proteins/genes involved with disease, drugs, and drug targets. The Venn diagrams represent the relationships between the drugome (the protein targets for currently approved drugs), the research drugome (the protein targets for drugs under research or in clinical trials), the diseaseome (the genes that have been associated with disease according to the Online Inheritance in Man [OMIM] database), and drugs (investigational and approved). The circular diagrams represent protein targets of drugs, and the elliptical diagram represents drugs. Approximately 500 protein drug targets are represented in the OMIM database, as depicted by the overlap of the drugome (red), research drugome (dark green), and diseaseome (pale green). Although not depicted by the black arrows in this figure, some approved drugs may have targets that are currently investigational.

in Man (OMIM) database and covers a tiny fraction of the proteome, which is estimated at approximately 100,000 proteins. The research drugome (the proteins targeted by approved drugs and/or drugs in research and clinical trials) overlaps with a portion (500 protein targets) of the diseaseome (Figure 1), and is focused in 15 to 20 therapeutic areas.[7,8] This means that about two-thirds of current drugs do not target a gene product that has been associated with disease.

The drug-discovery process is hindered by the cost and time of the research and development efforts necessary to take a drug from clinical trial to market: $1 billion and from 5 to 10 years, on average. Furthermore, many drugs fail in phase III trials due to lack of efficacy (no significant difference between the placebo and drug treatment or between drug and current treatment alternative). Considering the extremely high cost of new-drug development and the high drug attrition rate, it is evident that new approaches are required to improve and expedite the process. Since most current drug-discovery efforts are targeted toward complex diseases, understanding the pathophysiology at a systems level could aid in the design and development of drugs with higher success rates in clinical trials.

> *Since most current drug-discovery efforts are targeted toward complex diseases, understanding the pathophysiology at a systems level could aid in the design and development of drugs with higher success rates in clinical trials.*

The broadly defined field of systems biology has generated a large amount of accumulated knowledge on the biological processes involved in disease and drug action. In this review, we will provide an overview of how systems biology has been used to analyze the molecular mechanisms of human disease, current drug-target relationships, and mechanisms of targeted therapies, which is cumulatively termed "systems pharmacology."

ANALYZING MECHANISMS OF DISEASE: THROUGH REGULATORY NETWORKS

It is currently understood that single-gene disorders are rare and most genetic diseases are caused by the concerted action of many genetic abnormalities, along with other possible nongenetic influences such as environmental factors, hormone levels, age, weight, and diet. Consequently, the treatment of diseases with a single, "magic bullet" therapy has proven largely unsuccessful. The current understanding of pathophysiology is complicated by the large amount of information that has amassed over the years from biochemical and cell-biological studies, and a systems view is necessary to integrate all the factors that control the regulatory networks involved in orchestrating the physiological and pathophysiological responses.

In order to gain a systems-level, molecular perspective of disease, a network view of the relevant physiological function is essential.

> *In order to gain a systems-level, molecular perspective of disease, a network view of the relevant physiological function is essential.*

Each physiological process has an underlying signaling network of chemicals, hormones, native ligands, protein receptors, ions, enzymes, transcription factors, DNA with associated epigenetic mechanisms, and RNA, both regulatory and protein encoding. Together these components control the level of various cellular components to modulate biochemical reactions, electrical signals, force generation, transcription, and protein translation. These cellular processes can occur with different kinetics, simultaneously and/or at different times, and at varying levels of magnitude. Therefore, each physiological function or phenotype has a very complex network of signals controlling it. To represent each physiological process in a parsable network, each protein, ion, or chemical becomes a "node," and each interaction between 2 nodes (binding, chemical reaction, or other interaction) becomes an "edge" (Figure 2A, Table 1). Signaling networks can be represented as undirected graphs wherein interactions are specified but without hierarchy, or as directed graphs where the effect of 1 node on another is specified by the direction of the arrow. More specifically, directed graphs can have sign specification, such as positive for activation, negative for inhibition, or neutral interactions for scaffolds, which are represented differently by arrows, symbols, or colors. The hypothetical network in Figure 2A is

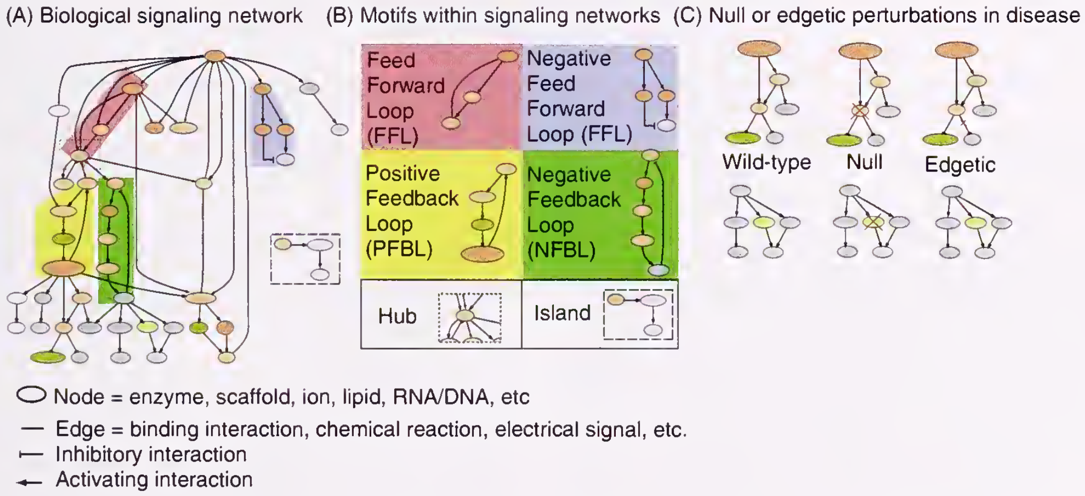

(A) Biological signaling network
(B) Motifs within signaling networks
(C) Null or edgetic perturbations in disease
Feed Forward Loop (FFL)
Negative Feed Forward Loop (FFL)
Positive Feedback Loop (PFBL)
Negative Feedback Loop (NFBL)
Hub
Island
Wild-type
Null
Edgetic
Node = enzyme, scaffold, ion, lipid, RNA/DNA, etc
Edge = binding interaction, chemical reaction, electrical signal, etc.
Inhibitory interaction
Activating interaction

Fig 2. Biological signaling networks and signaling motifs. (A) A hypothetical biological system is represented as a protein signaling network. Each protein or molecule in the network is a node (ovals) and each interaction is an edge. The network is directional and edges are given signs, with arrows representing positive or activating interactions and plungers representing negative or inhibitory interactions. Neutral interactions (scaffolding or anchoring interactions) are represented as lines. (B) The regulatory motifs within the protein signaling network are highlighted by color or dotted/dashed boxes. A feed-forward loop (FFL) is shown in pink, a negative FFL loop is shown in blue, a positive feedback loop is shown in yellow, and a negative feedback loop is shown in green. A hub is shown in a dotted box and an island is shown in a dashed box. (C) Mutations found in disease (left) are generally null (center) or edgetic (right). Null mutations, which are caused by gene truncations either by a lack of expression or expression of a severely shortened/compromised gene product, are represented as node removals in signaling networks. Edgetic mutations are caused by point mutations or small truncations leading to the expression of a gene product that has lost an interaction with another protein, and are represented edge removal(s).

Table 1. *Definitions of Key Terms in Systems Pharmacology.*

Term	Definition	Associated Figure
Drug target	A protein, typically an enzyme and/or transmembrane receptor, to which a drug binds and causes a change (positive or negative) in the action of.	1
Druggability	The likelihood of successfully designing a small molecule or biological molecule that can interfere with or promote the action of a given drug target.	
Genome	The overall genetic makeup of an organism; this includes coding and noncoding regions.	1
Proteome	The collection of proteins expressed by the genome.	1
Drugome	The collection of proteins that are targets of drugs.	1
Biological signaling network	Cell-signaling pathways represented as interactions between components within the system as edges and nodes. Such representations allow us to compute system properties.	2
Bipartite graph	A graph-theory term describing a network where the nodes are derived from 2 independent sets (examples cited in text are sets of drugs and drug targets or genes and diseases) and every edge connects a node from 1 group to the other group.	NA
Node	A representation of a component within a network that takes part in an interaction with another component in the network. Biological signaling network nodes can be enzymes, scaffolding proteins, ion channels, receptors, transcription factors, genes, ions, small molecules, or lipids.	2
Edge	A representation of an interaction between 2 components (nodes) within a biological signaling network. Edges represent binding reactions, gene expression, or enzymatic reactions such as post-translational modifications, molecular conversions (ie, ATP to cAMP).	2
Signaling motif	A pattern in the signaling flow within a network. Motifs include feedback loops, feed-forward loops, and bifan motifs.	2
Network redundancy	When >1 path in the signaling flow of a network leads to the same outcome.	NA

Abbreviations: ATP, adenosine triphosphate; cAMP, cyclic adenosine monophosphate; NA, not applicable.

directional, with activating interactions represented as arrows, inhibitory interactions as plungers, and neutral interactions as lines. In biological systems, signaling networks are not unidirectional[9]; instead, they are highly clustered. Certain nodes, termed "hubs" (Figure 2B), are highly connected,[10] thus allowing many more components to be connected to other subnetworks. On the other hand, a system can contain "islands" (Figure 2B), sets of nodes not connected to the overall network. Several types of motifs exist within regulatory networks, such as negative feedback loops, positive feedback loops, and feed-forward loops (Figure 2B).[9] These characteristics, as well as the high clustering coefficient, are attributed to the scale-free (a connectivity distribution that follows a power law[11]) and redundant properties of biological networks. These properties also allow for network perturbation without complete loss of function. Thus, a single imbalance in the regulatory network of a biological system, such as an increased protein level or biochemical reaction rate, does not necessarily lead to a disease state. Instead, within the regulatory network of a disease system, the signaling networks underlying the pathophysiology are likely to be persistently perturbed at more than 1 point (edge or node).

The genomic irregularities associated with disease have been studied, through genome-wide association and positional cloning studies, which have become more efficient with the sequencing of the human genome. More recently, numerous genomic, short hairpin/small interfering RNA (shRNA/siRNA), proteomic, and phosphoproteomic studies have focused on the gene-expression signatures and protein-signaling pathways associated with a particular disease (Table 2).[12–17] The plethora of data on the genetic mutations and variations produced by studies such as these has been compiled into several databases, such as the OMIM database and the Catalogue of Somatic Mutations in Cancer (COSMIC) database. These global databases help to identify genes, proteins, and other cellular

Table 2. *Systems Biology Studies That Elucidate Mechanisms of Pathophysiology and Drug Action.*

Methodology	Description	Examples
Genomics and proteomics	Large-scale experimental analysis of gene expression, genomic or epigenetic variations; post-translational changes resulting from varying conditions; includes phosphoproteomics and chemical proteomics. Cells or tissue samples are subject to a condition (eg, drug treatment, stress, diseased versus wild type/normal) and the response is analyzed compared with a control. Expression level, genetic variation, phosphorylation status, and other post-translational modifications can be monitored.	16,46–48 21,49,50
Large-scale molecular screening and drug structure-target analyses	Experimental and computational methods uncover unknown targets and/or promiscuity of drugs. These analyses can allow for further characterization the targets that are inhibited by common drugs, or vice versa.	6,23,42,49,51,52,53
shRNA/siRNA screens	A high-throughput version of silencing a single gene and analyzing the effect in cells. Resulting information can explain/predict the function of a protein in disease or the effect of inhibiting the protein as a means of drug-target validation.	54–57 58,59
Network analysis and computational methods	The identification and understanding of networks surrounding a drug target, often using published data from high-throughput studies (such as those listed above) or databases.	8,9,20,22,60,61

Abbreviations: shRNA, short hairpin RNA; SILAC, stable isotope labeling by amino acids in cell culture; siRNA, small interfering RNA.

components related to the origin or progression of a specific disease. These components can be used to generate lists of "seed nodes" that enable us to computationally construct various types of networks by integrating knowledge about gene-disease associations and/or protein interactions reported in biomedical literature. Through information about the properties of disease genes as nodes in a network, previously unrelated relationships between disease pathways and cellular functions can be identified, which can serve as a first step in identifying drug targets.[18–20]

An interactive, computable database, the Connectivity Map,[21] finds correlations between gene-expression signatures and the sets of proteins involved with the action of a class of drugs or with a particular disease. The initial Connectivity Map project analyzed the genome-wide messenger RNA (mRNA) expression profiles of 1 cell type after treatment with 164 small molecules and perturbagens (a subset was tested in 3 other cell lines). Instead of using traditional hierarchical clustering for analyzing the genetic data, the authors developed and adopted a method termed "gene set enrichment analysis," a nonparametric, rank-based pattern-matching strategy,[21] with the idea that regardless of the cell type, the gene-expression data specific to perturbation by a small molecule could be detected. The gene-expression signatures could be queried to search for commonalities among expression signatures for seemingly unrelated drugs or drug classes. This tool aids in the empirical understanding of current drug action and the discovery and characterization of new drug entities.[22]

In another example, Goh *et al.*[19] examined gene-disorder association data in 2 networks: (1) the human-disease network, which treats human diseases as nodes, connected based on the associated genes they have in common, and (2) the disease-gene network, which connects disease genes according to whether they are associated with the same disorder. The data were used to construct a bipartite interaction network using 1284 known disorders and 1177 known disease genes from the OMIM database. The diseases were further broken down into 22 classes. Both networks were highly connected and had a "giant component," a hub-like area where many nodes were connected to one another. In the human disease network, cancer and related disorders and neurological disorders were in highly connected, heterogeneous hubs, whereas metabolic, skeletal, and other disorders were in lesser-connected "islands." With a wide distribution of gene associations, most diseases associate with a few genes, whereas some disorders associated with a much higher number of genes (30 to 50). There were numerous islands, more notably in the disease-gene network. It was previously demonstrated that disease-related genes are hubs or have a higher number of interactions than nondisease proteins.

The majority of the gene list used in the human disease and disease gene networks contained nonessential genes ("essential" defined as being embryonic/postnatal lethal).[19] It was observed that with essential genes excluded from the disease-gene network, the connectivity greatly diminishes, indicating that the higher connectivity was observed solely due to the high connectivity inherent to the essential genes. Furthermore, the expression patterns of disease genes were asynchronous with genes essential for cellular functions, and the disease genes had little overlap with housekeeping genes. This indicated that disease genes are often peripheral, with 1 important exception: somatic mutations found in cancer, which did not fit the analysis described above.

In another analysis of the properties of disease genes from a systems perspective, Zhong *et al.*[20] showed that genetic mutations that perturb networks at an edge (edgetic: mutated nearly full-length or point-mutated full-length gene products) rather than at a node (null: truncated gene products) are more heavily responsible for the phenotypic manifestation of the diseases.[20] Edgetic perturbations are more commonly found in mutations that span multiple diseases and are caused by in-frame mutations and the expression of a nearly full-length gene product, which lacks an interaction with a neighboring node (Figure 2C). In contrast, node removal or a null mutation includes severely truncated gene products or other mutations that destabilize protein structure and equate to the removal of the protein with respect to its function and interactions (Figure 2C). Edgetic perturbations have potential consequences at more than 1 node in the network. Furthermore, the authors cited examples where different mutations of the same gene product affect different edges, and each edgetic mutation is associated with a different disorder.

ANALYZING DRUG-TARGET RELATIONSHIPS WITH REGULATORY NETWORKS

Systematically organizing relationships between diseases, drugs, and proteins into a map such as a signaling network enables statistical inferences about less-obvious relationships between them. Global analyses of FDA-approved drugs have found that, unlike ubiquitously expressed essential genes, drug targets tend to be expressed in specific tissues but interact with many other proteins in the cellular network while being independently regulated.[8] Commonalities of signaling mechanisms among existing drug targets create a basis for identifying the network properties that define a potentially good drug target. Statistical metrics from network analyses, such as a centrality measure, which quantifies the relative importance of a protein in communicating between different modules within a network, have been suggested for identifying nodes (proteins in a network) that have attractive properties as potential drug targets.[23]

Ma'ayan *et al.* constructed a bipartite graph of 1052 FDA-approved drugs interacting with 485 targets contained 179 islands, using data from DrugBank.[24,25] Most of these islands are made up of 10 to 30 interacting cellular components. Drugs that target GPCRs comprise a single large island of 481 components.[26] A majority of targets of currently approved drugs are GPCRs (~25%),[5] which govern a diversity of physiological processes such as cardiac contractility, acid secretion, and airway constriction.

In a study by Yildirim *et al.*,[8] a bipartite graph of drugs and their targets listed in DrugBank was generated. Analysis of this drug-target network showed that the majority (62%) of currently approved drugs target membrane proteins, most likely the bulk of which are GPCRs, as indicated by the results of study by Ma'ayan *et al.*[26] In the analysis of the distribution of the targets of research and current drugs, the drug-target identity distribution became more diverse, with cytoplasmic targets doubling (6% to 12%). This result is indicative that drug-discovery technologies have improved to target nonmembrane proteins and the emergence of targeted therapies.

TARGETED THERAPIES: ADVANTAGES AND PITFALLS

Targeted therapies or signal transduction inhibitors make up the bulk of the current drug pipeline. Of all drug targets, the majority (25%) target GPCRs, followed by kinases (10%).[5] The shift from GPCR drugs to other receptor classes and protein kinase–targeted drugs (>100 kinase inhibitors are in development and 15 are currently approved as cancer therapies, Table 3) is representative of the cumulative knowledge of the mechanisms underlying disease. The obvious advantage of targeted therapies is their mechanistic derivation, targeted to the specific signaling molecule (typically a kinase) that exhibits aberrant behavior in disease. Many targeted therapies target a new area and the risk of failure is higher if the drug has a novel mechanism of

Table 3. *Currently Approved Targeted Therapies for the Treatment of Cancer.*[62–65]

Therapy (Trade Name) Company	Target (Primary in bold)	Disease
Imatinib (Gleevec) Novartis	**Abl**, c-Kit, PDGFR, CSF1R, NTRK1, DDR1	CML, ALL, CEL, GIST
Dasatinib Sprycel	**Src, Abl**, c-Kit, PDGFR	CML (imatinib resistant), ALL
Nilotonib (Tasigna) Novartis	**Bcr-Abl, Abl**, PDGFR, c-KIT, DDR1	CML (imatinib resistant)
Erlotinib (Tarceva) OSI Genentech/Roche	**EGFR, ErbB2**	NSCLC, pancreatic cancer
Gefitinib (Iressa) AstraZeneca	**EGFR**, ErbB2, ErbB4	NSCLC
Sorafenib (Nexavar) Bayer/Onyx	**B-Raf**, c-KIT, PDGFR, VEGFR	Renal cancer
Sunitinib (Sutent) Pfizer	**c-KIT, PDGFR, VEGFR, KIT, RET, CSF1R**	Renal cancer, GIST (imatinib resistant)
Lapatinib (Tykerb) GlaxoSmithKline	**ErbB2**	Breast cancer
Pazopanib (Votrient) GlaxoSmithKline	**VEGFR, PDGFR, c-Kit**	Renal cancer
Trastuzumab (Herceptin) Genentech	**ErbB2**	Breast cancer

Abbreviations: ALL, acute lymphoblastic leukemia; CEL, chronic eosinophilic leukemia; CML, chronic myelogenous leukemia; GIST, gastrointestinal stromal tumor; NSCLC, non-small cell lung cancer.

action.[27] Other disadvantages lie in the difficulty of achieving high specificity and the resulting off-target effects and primary and acquired mechanisms of resistance to these therapies. Complex signaling mechanisms and motifs often play a role in resistance, underscoring the importance of a systems-pharmacology understanding in the target choice and drug design.

> *Complex signaling mechanisms and motifs often play a role in resistance, underscoring the importance of a systems-pharmacology understanding in the target choice and drug design.*

The bulk of targeted therapies is used to treat cancer, which this discussion will focus on. The first example of a targeted therapy in cancer was tamoxifen, which specifically inhibits the estrogen receptor in estrogen receptor–positive breast cancers.[28] The first successful kinase inhibitor was the BCR-ABL inhibitor imatinib, which is approved for treating chronic myelogenous leukemia. Other examples (Table 2) include human epidermal growth factor receptor family (EGFR and HER-2/ErbB2) inhibitors such as monoclonal antibody therapies trastuzumab, cetuximab, and panitumumab, and the tyrosine kinase inhibitors erlotinib, gefitinib, and lapatinib.

By design, these therapies are administered to select patient populations that exhibit a specific mutation or heightened expression of the protein target. For example, prior to receiving trastuzumab, patients must test positive for elevated ErbB2 expression[29,30]; and to be enrolled in a cetuximab combination clinical trial for pancreatic cancer, biopsies must show heightened EGFR expression.[31] On the other hand, patients are also tested for primary mechanisms of resistance to these drugs, such as K-Ras mutations mitigating the response to EGFR inhibitors. Because these drugs target a specific signaling molecule, acquired resistance is a significant problem. For example, the T790M mutation in EGFR increases the affinity of adenosine triphosphate (ATP) for the EFGR active site but ablates the affinity of the EGFR kinase inhibitor gefitinib. Also, various mutations in mitogen-activated/extracellular signal-regulated kinase kinase (MEK) were exhibited by tumors which relapsed following treatment with the MEK inhibitor AZD6244, which is currently in clinical trials.[32]

Other types of acquired resistance are caused by compensatory mechanisms, which are distinct from those encountered with nonselective chemotherapies which result from increase in the levels of the drug transporters like the ATP-binding cassette protein family. Instead, the resistance mechanisms involve nodes within the signaling network of the targeted protein. One example of resistance to EGFR therapy is the maintenance of phosphoinositide 3-kinase (PI3K) signaling through a variety of mechanisms, including loss of phosphatase and tensin homolog (PTEN) (a phosphatase and negative regulator of V-akt murine thymoma viral oncogene homolog [AKT]), ErbB3, or methionine (Met) amplification, which allows for tumor cell survival despite EGFR inhibition.[33] Additionally, in gefitinib-resistant and cetuximab-resistant cells, the HER signaling switches from EGFR-containing heterodimers and homodimers to primarily HER2-Her3 signaling dimers, which more potently activates the PI3K/AKT pathway.[32,34]

In targeting a specific component in a complex signaling pathway, there is potential to disrupt all of the intrinsic signaling mechanisms associated with that particular node, which can have counterintuitive results.

> *In targeting a specific component in a complex signaling pathway, there is potential to disrupt all of the intrinsic signaling mechanisms associated with that particular node, which can have counterintuitive results.*

For example, due to prevalence of mutations upstream of extracellular signal-regulated kinase (ERK) in cancers, such as mutant Ras and B-Raf, several B-Raf and MEK inhibitors are currently approved or in development. Due to the complexity of the signaling pathways, targeted inhibition of these molecules in cancer cells can cause unwanted signal amplification or bypass the inhibition within the targeted network. For instance, overactive B-Raf will hyperactivate the MEK–mitogen-activated protein kinase (MEK-MAPK) pathway, similar to oncogenic Ras, and inhibition of MEK should normalize this aberrancy, whether it originates with B-Raf or Ras mutations (Figure 3). However, MEK inhibition caused an increase in PI3K/AKT activation, and, vice versa, PI3K/AKT inhibition caused an increase in MAPK activation (Figure 3B).[35,36] These paradoxical results are caused by a negative feedback loop in receptor tyrosine kinase (RTK) signaling that is mediated by TSC2-P70S6K and controlled by AKT and ERK. Hence, when MEK/ERK or PI3K/AKT is inhibited, a negative feedback loop is partially

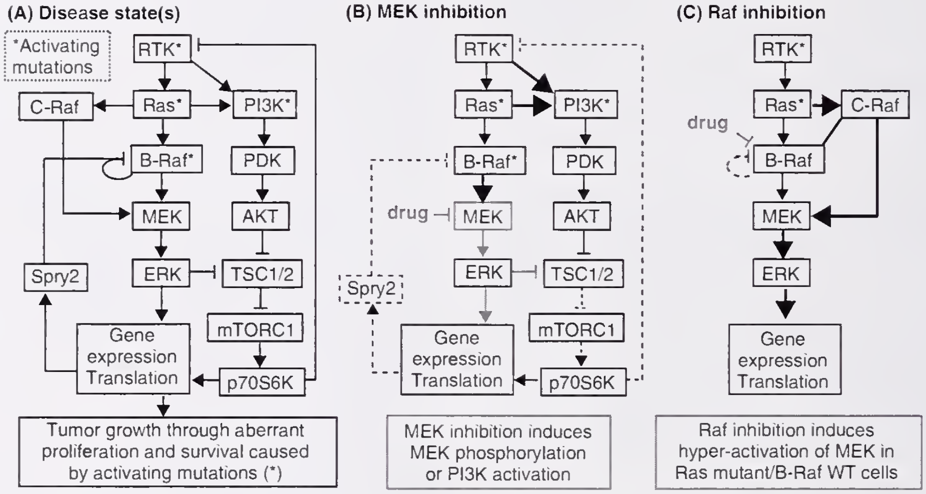

(A) Disease state(s)
*Activating mutations
RTK*
C-Raf
Ras*
PI3K*
B-Raf*
PDK
MEK
AKT
Spry2
ERK
TSC1/2
mTORC1
Gene expression Translation
p70S6K
Tumor growth through aberrant proliferation and survival caused by activating mutations (*)
(B) MEK inhibition
RTK*
Ras*
PI3K*
B-Raf*
PDK
drug
MEK
AKT
ERK
TSC1/2
Spry2
mTORC1
Gene expression Translation
p70S6K
MEK inhibition induces MEK phosphorylation or PI3K activation
(C) Raf inhibition
RTK*
Ras*
C-Raf
drug
B-Raf
MEK
ERK
Gene expression Translation
Raf inhibition induces hyper-activation of MEK in Ras mutant/B-Raf WT cells

Fig 3. Signaling intricacies determine therapeutic outcome in complex disease. (A) In complex diseases such as cancer, 1 or more activating mutations cause cellular transformation, enhanced cell growth, proliferation, and survival, leading to tumor formation and maintenance. Drugs have been developed to inhibit hyperactive signaling pathway(s) caused by the mutations in cancer, such as MEK and B-Raf inhibitors. (B) When cells are treated with a MEK inhibitor (in red), disruption of negative feedback loops (dashed lines) causes counterintuitive increases in signaling upstream of MEK (large bold arrows), such as phosphorylation of MEK by B-Raf and activation of PI3K indirectly by the RTK. (C) When B-Raf WT/Ras mutant cells are treated with B-Raf inhibitors (in red), the autoinhibitory property of B-Raf is reduced and its affinity for C-Raf is increased. This results in recruitment of B-Raf/C-Raf to Ras at the membrane, where C-Raf is activated, causing hyperactivation of MEK. When B-Raf mutant cells are treated with B-Raf inhibitors, this effect is not seen and MEK remains inhibited. **Abbreviations:** AKT, V-akt murine thymoma viral oncogene homolog; ATP, adenosine triphosphate; B-Raf, V-raf murine sarcoma viral oncogene homolog B1; C-Raf, V-raf-1 murine leukemia viral oncogene homolog 1; ERK, extracellular signal-regulated kinase; GPCR, G-protein-coupled receptor; mTORC1, mammalian target of rapamycin complex 1; MEK, mitogen-activated extracellular signal-regulated kinase kinase; Met, methionine; mRNA, messenger RNA; PDK, 3-phosphoinositide dependent protein kinase; PI3K, phosphoinositide 3-kinase; PTEN, phosphatase and tensin homolog; RTK, receptor tyrosine kinase; shRNA, short hairpin RNA; siRNA, small interfering RNA; Spry2, Sprouty homolog 2 (drosophila); TSC1/2, tuberous sclerosis complex tumor suppressor 1–2; WT, wild type.

shut down, causing signal amplification in the noninhibited arm of the pathway (Figure 3B).

In another example, treatment with the MEK inhibitor (AZD6244) reduced proliferation cells that harbor the activating B-Raf V600E mutation.[37] However, in cells harboring activating K-Ras mutations with wild-type B-Raf, AZD62444 treatment caused accumulation of phospho-MEK, partially due to ablation of a negative feedback loop between MEK and B-Raf, mediated by Sprouty2 (Figure 3B).[38] This negative feedback loop is disrupted by the B-Raf V600E mutation, due to reduced affinity between mutant B-Raf and Sprouty2. Hence, response to AZD62444 is observed in cells harboring the V600E mutation.

Two recent, pertinent studies describe how B-Raf inhibition drives Ras-mediated B-Raf binding to and activation of C-Raf, followed by MEK-ERK activation (Figure 3C).[39,40] In some B-Raf/Ras wild type (WT) cell lines and B-Raf WT/Ras-mutant cell lines, Raf-inhibitor treatment caused an increase in tumor cell proliferation. B-Raf inhibitor(s)' binding of the B-Raf ATP pocket potentiated the formation of B-Raf/C-Raf heterodimers at the membrane, in a Ras-dependent manner. Increased heterodimer formation correlated with increased C-Raf activity and increased MEK activation. MEK activation did not occur in drugs that inhibit both B-Raf and C-Raf.[39] Another study on AKT inhibitors produced similar results. When AKT was specifically inhibited, a pool of hyperphosphorylated AKT occurred, due to the "priming"caused by inhibitor binding the ATP pocket of AKT.[41]

These scenarios exemplify that network-signaling motifs play a role in drug action, highlighting the important role that the intricacies of the signaling pathways play in governing the effects of these targeted therapies.

> *These scenarios exemplify that network-signaling motifs play a role in drug action, highlighting the important role that the intricacies of the signaling pathways play in governing the effects of these targeted therapies.*

A challenge in applying systems-pharmacology methods in translational medicine is the identification of complex signaling mechanisms early in the drug-discovery process. A first and necessary step to address this challenge in the clinic is obtaining as much data as possible on pertinent mutations and genomic variations, especially with complex diseases in which variable and multiple genetic factors lead to a diseased state. While a set of parameters does not exist to classify the gene products or interactions that will play a critical role in disease or drug action, establishing a systems perspective can lead to the detection of unforeseen effects of modulating a single component in a biological signaling network. In the early stages of considering a drug target, the following steps can be taken to enable a systems perspective and all the possibility of uncovering unforeseen consequences of modulating the activity of the drug target: (1) Using the literature and databases, identify all components that interact with the drug target of interest and the signaling pathways/networks with which these components are involved. (2) Construct a directed graph of all of these components (Figure 2), including all direct interactions and ones that lead to physiologically significant events (ie, ones that lead to a major result such as transcription, cell survival/death, motility, or others). (3) Identify redundant pathways and signaling motifs such as feedback loops, which may cause signal amplification, compensation, or other effects in the specific system of interest. (4) Establish a cell-based model to test these scenarios and take into account the results when considering the drug target and combining or prescribing drugs in a given genetic background.

COMPUTATIONAL METHODOLOGIES IN DRUG DISCOVERY

Often, drugs bind to several targets that are seemingly unrelated except for the target-binding pattern of the drug. Recent work has generated a network connecting drugs on the basis of their structural similarity and similarity of side-effect profiles.[42] This method has been proven effective at identifying groups of drugs that share common targets. When drugs with predicted shared targets did not have known targets in common, the predictions provided a framework for testing these drugs for binding against the predicted targets. Hence, this approach led to the identification of new targets for existing drugs, which serve as new therapeutic targets or explain the cause(s) of an adverse effect.

There have been recent developments in "chemo-centric" methodologies to systematically compare the structures of drugs with known targets, as a means of predicting unknown target binding.

> *There have been recent developments in "chemo-centric" methodologies to systematically compare the structures of drugs with known targets, as a means of predicting unknown target binding.*

The idea of searching a database of structures for a similar structure has been around a long time in the field of chemistry. For decades, chemists have been using structure-based literature searches for reaction methods, mechanisms, chemical information, and various analyses.[43] New approaches allow databases of structural and pharmacological data to be parsed using chemical structure, provide information on the global properties of compound libraries with respect to their biological properties, and predict binding interactions.

A major feat in establishing this methodology was the assembly of a large database of chemical entities, of which ~275,000 are curated with structure-activity relationship data (target identity, IC_{50}, EC_{50}, K_i, and K_D values) from >2800 targets.[44] The human subset of targets (700) and the respective drugs they bind was used to create a map of the protein targets (nodes) and how they are connected in chemical space (1 molecule binding 2 proteins creates an edge). The promiscuity of each target was analyzed, with aminergic GPCRs, cytochrome P450, and protein kinases ranking the highest. In a similar analysis, it was found that compounds generally targeted proteins from the same gene families. The method was also used to predict pharmacology of compounds and analyze how the physical properties of drugs differ within each target family. Unfortunately, the database is not publicly available and is proprietary to Pfizer, where the study took place,[44] but it represents a cumulative search tool, a novel method of searching for new leads in the drug-discovery process and improving the physical properties of lead compounds.

Another approach, coined the Similarity Ensemble Approach (SEA), classifies ligands, with known receptor interactions, by their chemical similarity using the Tanimoto coefficient of chemical similarity (Tc), a fragment-based algorithm.[45] Unrelated structures are parsed with these codes to predict with which protein targets they might interact. Comparing drug-target interactions by way of protein structure comparisons is much more difficult than by way of drug structure comparisons, due to the sheer difference in size and structural complexity. Central to this methodology is the idea that small molecules with similar structures will display similar target-binding properties. A recent study used this method and predicted that Paxil and Prozac (serotonin reuptake inhibitors) are also β-blockers.[23] The binding of these drugs to the β-adrenergic receptor was experimentally verified. The authors speculated that some observed adverse effects of the drugs, such as increased heart rate and sexual dysfunction, could be attributed to the off-target binding.

CONCLUSION

Basic cell biological research continues to delineate the functions of each protein in the proteome as well as metabolites in the metabolome to construct the "interactome" (how all cellular components interact) in both physiological and pathophysiological systems. The construction of the interactomes for specific pathophysiology can be quite useful for uncovering potential new drug targets. The combination of such physiological analyses at a systems level with informatics approaches based on drug structure and structure-function relationships offers a powerful new approach for drug discovery in the preclinical phase. The drug-discovery platforms have shifted toward targeted therapies, and drug-design technologies have greatly advanced to include biological therapies (antibodies and peptides), antisense therapies, and other methods to modulate protein-protein interactions, expression levels, or other interactions. However, more potential drug targets and drug-discovery technologies also means more possible efficacy failure and adverse effects that will set back the efforts to move drugs to the clinics. To enhance the rate at which new drugs are brought into the clinic, one will need to understand pharmacokinetic and pharmacodynamics of new drugs and use this knowledge in clinical-trial design. A systems-biology perspective and advanced computation methodologies provide an integrated basis for the understanding of the complex mechanisms of disease, targeted therapy action, and design combination therapies, and thus influence clinical-trial design to rapidly test new drugs in the clinic.

ACKNOWLEDGMENTS

We thank Seth Berger for help with the drug-target analyses associated with Figure 1 and Sherry Jenkins for valuable feedback. This research is supported by the Systems Biology Center Grant P50-GM 071 558.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Kenakin T. Principles: receptor theory in pharmacology. *Trends Pharmacol Sci* 2004; 25: 186–192.
2. Black JW, Leff P. Operational models of pharmacological agonism. *Proc R Soc Lond B Biol Sci* 1983; 220: 141–162.
3. Maehle AH, Prüll CR, Halliwell RF: The emergence of the drug receptor theory. *Nat Rev Drug Discov* 2002; 1: 637–641.
4. Colquhoun D. The quantitative analysis of drug-receptor interactions: a short history. *Trends Pharmacol Sci* 2006; 27: 149–157.
5. Hopkins AL, Groom CR. The druggable genome. *Nat Rev Drug Discov* 2002; 1: 727–730.
6. Russ AP, Lampel S. The druggable genome: an update. *Drug Discov Today* 2005; 10: 1607–1610.
7. Gao Z, Li H, Zhang H, et al. PDTD: a web-accessible protein database for drug target identification. *BMC Bioinformatics* 2008; 9: 104.
8. Yildirim MA, Goh KI, Cusick ME, et al. Drug-target network. *Nat Biotechnol* 2007; 25: 1119–1126.
9. Ma'ayan A, Blitzer RD, Iyengar R. Toward predictive models of mammalian cells. *Annu Rev Biophys Biomol Struct* 2005; 34: 319–349.
10. Ravasz E, Somera AL, Mongru DA, et al. Hierarchical organization of modularity in metabolic networks. *Science* 2002; 297: 1551–1555.
11. Jeong H, Tombor B, Albert R, et al. The large-scale organization of metabolic networks. *Nature* 2000; 407: 651–654.
12. Ideker T, Sharan R. Protein networks in disease. *Genome Res* 2008; 18: 644–652.
13. Hua S, Kallen CB, Dhar R, et al. Genomic analysis of estrogen cascade reveals histone variant H2A.Z associated with breast cancer progression. *Mol Syst Biol* 2008; 4: 188.
14. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005; 6: 95–108.
15. Speer R, Wulfkuhle J, Espina V, et al. Molecular network analysis using reverse phase protein microarrays for patient tailored therapy. *Adv Exp Med Biol* 2008; 610: 177–186.
16. Macek B, Mann M, Olsen JV. Global and site-specific quantitative phosphoproteomics: principles and applications. *Annu Rev Pharmacol Toxicol* 2009; 49: 199–221.
17. Jimenez-Sanchez G, Childs B, Valle D. The effect of Mendelian disease on human health. In: Scriver CR, Beaudet AR, Sly W, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York, NY: McGraw-Hill; 2001; 167–174.
18. Feldman I, Rzhetsky A, Vitkup D. Network properties of genes harboring inherited disease mutations. *Proc Natl Acad Sci U S A* 2008; 105: 4323–4328.
19. Goh KI, Cusick ME, Valle D, et al. The human disease network. *Proc Natl Acad Sci U S A* 2007; 104: 8685–8690.
20. Zhong Q, Simonis N, Li QR, et al. Edgetic perturbation models of human inherited disorders. *Mol Syst Biol* 2009; 5: 321.
21. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006; 313: 1929–1935.
22. Lamb J: The Connectivity Map: a new tool for biomedical research. *Nat Rev Cancer* 2007; 7: 54–60.
23. Keiser MJ, Setola V, Irwin JJ, et al. Predicting new molecular targets for known drugs. *Nature* 2009; 462: 175–181.
24. Wishart DS, Knox C, Guo AC, et al. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res* 2008; 36(Database issue): D901–D906.
25. Wishart DS, Knox C, Guo AC, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* 2006; 34(Database issue): D668–D672.
26. Ma'ayan A, Jenkins SL, Goldfarb J, Iyengar R. Network analysis of FDA approved drugs and their targets. *Mt Sinai J Med* 2007; 74: 27–32.
27. Gordian M, Singh N, Zemmel R, Elias T. Why products fail in phase III. *In Vivo* 2006; 4: 49–54.
28. Jordan VC. Tamoxifen (ICI46,474) as a targeted therapy to treat and prevent breast cancer. *Br J Pharmacol* 2006; 147(suppl 1): S269–S276.
29. Harris L, Fritsche H, Mennel R, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol* 2007; 2: 5287–5312.
30. Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005; 353: 1673–1684.
31. Cascinu S, Berardi R, Labianca R, et al. Cetuximab plus gemcitabine and cisplatin compared with gemcitabine and cisplatin alone in patients with advanced pancreatic cancer: a randomised, multicentre, phase II trial. *Lancet Oncol* 2008; 9: 39–44.
32. Emery CM, Vijayendran KG, Zipser MC, et al. MEK1 mutations confer resistance to MEK and B-RAF inhibition. *Proc Natl Acad Sci U S A* 2009; 106: 20411–20416.
33. Klein S, Levitzki A. Targeting the EGFR and the PKB pathway in cancer. *Curr Opin Cell Biol* 2009; 21: 185–193.
34. Jain A, Penuel E, Mink S, et al. HER kinase axis receptor dimer partner switching occurs in response to EGFR tyrosine kinase inhibition despite failure to block cellular proliferation. *Cancer Res* 2010; 70: 1989–1999.
35. Sos ML, Fischer S, Ullrich R, et al. Identifying genotype-dependent efficacy of single and combined PI3K- and MAPK-pathway inhibition in cancer. *Proc Natl Acad Sci U S A* 2009; 106: 18351–18356.
36. Carracedo A, Ma L, Teruya-Feldstein J, et al. Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. *J Clin Invest* 2008; 118: 3065–3074.
37. Friday BB, Adjei AA. Advances in targeting the Ras/Raf/MEK/Erk mitogen-activated protein kinase

cascade with MEK inhibitors for cancer therapy. *Clin Cancer Res* 2008; 14: 342–346.

38. Friday BB, Yu C, Dy GK, et al. BRAF V600E disrupts AZD6244-induced abrogation of negative feedback pathways between extracellular signal-regulated kinase and Raf proteins. *Cancer Res* 2008; 68: 6145–6153.
39. Heidorn SJ, Milagre C, Whittaker S, et al. Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. *Cell* 2010; 140: 209–221.
40. Hatzivassiliou G, Song K, Yen I, et al. RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth. *Nature* 2010; 464: 431–435.
41. Okuzumi T, Fiedler D, Zhang C, et al. Inhibitor hijacking of Akt activation. *Nat Chem Biol* 2009; 5: 484–493.
42. Campillos M, Kuhn M, Gavin AC, et al. Drug target identification using side-effect similarity. *Science* 2008; 321: 263–266.
43. Willett P. Searching techniques for databases of two- and three-dimensional chemical structures. *J Med Chem* 2005; 48: 4183–4199.
44. Paolini GV, Shapland RH, van Hoorn WP, et al. Global mapping of pharmacological space. *Nat Biotechnol* 2006; 24: 805–815.
45. Keiser MJ, Roth BL, Armbruster BN, et al. Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* 2007; 25: 197–206.
46. Manning G, Whyte DB, Martinez R, et al. The protein kinase complement of the human genome. *Science* 2002; 298: 1912–1934.
47. Oppermann FS, Gnad F, Olsen JV, et al. Large-scale proteomics analysis of the human kinome. *Mol Cell Proteomics* 2009; 8: 1751–1764.
48. VanMeter A, Signore M, Pierobon M, et al. Reverse-phase protein microarrays: application to biomarker discovery and translational medicine. *Expert Rev Mol Diagn* 2007; 7: 625–633.
49. Bamborough P, Drewry D, Harper G, et al. Assessment of chemical coverage of kinome space and its implications for kinase drug discovery. *J Med Chem* 2008; 51: 7898–7914.
50. Ong SE, Schenone M, Margolin AA, et al. Identifying the proteins to which small-molecule probes and drugs bind in cells. *Proc Natl Acad Sci U S A* 2009; 106: 4617–4622.
51. Frye SV. Structure-activity relationship homology (SARAH): a conceptual framework for drug discovery in the genomic era. *Chem Biol* 1999; 6: R3–R7.
52. Fuller JC, Burgoyne NJ, Jackson RM. Predicting druggable binding sites at the protein-protein interface. *Drug Discov Today* 2009; 14: 155–161.
53. Lovering F, Bikker J, Humblet C. Escape from flatland: increasing saturation as an approach to improving clinical success. *J Med Chem* 2009; 52: 6752–6756.
54. Root DE, Hacohen N, Hahn WC, et al. Genome-scale loss-of-function screening with a lentiviral RNAi library. *Nat Methods* 2006; 3: 715–719.
55. Azorsa DO, Robeson RH, Frost D, et al. High-content siRNA screening of the kinome identifies kinases involved in Alzheimer's disease-related tau hyperphosphorylation. *BMC Genomics* 2010; 11: 25.
56. Gumireddy K, Li A, Gimotty PA, et al. KLF17 is a negative regulator of epithelial-mesenchymal transition and metastasis in breast cancer. *Nat Cell Biol* 2009; 11: 1297–1304.
57. Meacham CE, Ho EE, Dubrovsky E, et al. In vivo RNAi screening identifies regulators of actin dynamics as key determinants of lymphoma progression. *Nat Genet* 2009; 41: 1133–1137.
58. Chang K, Elledge SJ, Hannon GJ. Lessons from Nature: microRNA-based shRNA libraries. *Nat Methods* 2006; 3: 707–714.
59. Gaither A, Iourgenko V. RNA interference technologies and their use in cancer research. *Curr Opin Oncol* 2007; 19: 50–54.
60. Berger SI, Iyengar R. Network analyses in systems pharmacology. *Bioinformatics* 2009; 25: 2466–2472.
61. Iyengar R. Why we need quantitative dynamic models. *Sci Signal* 2009; 2: eg3.
62. DrugBank. http://www.drugbank.ca.
63. National Cancer Institute. Targeted cancer therapies. http://www.cancer.gov/cancertopics/factsheet/Therapy/targeted.
64. Janne PA, Gray N, Settleman J. Factors underlying sensitivity of cancers to small-molecule kinase inhibitors. *Nat Rev Drug Discov* 2009; 8: 709–723.
65. US Food and Drug Administration. http://www.fda.gov.

Protein Kinase Target Discovery From Genome-Wide Messenger RNA Expression Profiling

Avi Ma'ayan[1,2,3] and John C. He[1,4,5]

[1]Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY
[2]Systems Biology Center New York, Mount Sinai School of Medicine, New York, NY
[3]Experimental Therapeutics Institute, Mount Sinai School of Medicine, New York, NY
[4]Department of Medicine, Mount Sinai School of Medicine, New York, NY
[5]James J. Peters VA Medical Center, Bronx, New York

OUTLINE

DETECTING CHANGES IN MESSENGER RNA EXPRESSION
PROMOTER ANALYSIS AND CHROMATIN IMMUNOPRECIPITATION ENRICHMENT ANALYSIS
CONNECTING IDENTIFIED TRANSCRIPTION FACTORS
VARIATIONS IN WORKFLOW
CONCLUSION
ACKNOWLEDGEMENTS

ABSTRACT

Genome-wide messenger RNA profiling provides a snapshot of the global state of the cell under different experimental conditions such as diseased versus normal cellular states. However, because measurements are in the form of quantitative changes in messenger RNA levels, such experimental data does not provide direct understanding of the regulatory molecular mechanisms responsible for the observed changes. Identifying potential cell signaling regulatory mechanisms responsible for changes in gene expression under different experimental conditions or in different tissues has been the focus of many computational systems biology studies. Most popular approaches include promoter analysis, gene ontology, or pathway enrichment analysis, as well as reverse engineering of networks from messenger RNA expression data. Here we present a rational approach for identifying and ranking protein kinases that are likely responsible for observed changes in gene expression. By combining promoter analysis; data from various chromatin immunoprecipitation studies such as chromatin immunoprecipitation sequencing, chromatin immunoprecipitation coupled with paired-end ditag, and chromatin immunoprecipitation-on-chip; protein-protein interactions; and kinase-protein phosphorylation reactions collected from the literature, we can identify and rank candidate protein kinases for knock-down, or other types of functional validations, based on genome-wide changes in gene expression. We describe how protein kinase candidate identification and ranking can be made robust by cross-validation with phosphoproteomics data as well as through a literature-based text-mining approach. In conclusion, data integration can produce robust candidate rankings for understanding cell regulation through identification of protein kinases responsible for gene expression changes, and thus rapidly advancing drug target discovery and unraveling drug mechanisms of action. ***Mt Sinai J Med 77:345–349, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: cell signaling, ChIP-X enrichment analysis, computational systems biology, mRNA profiling, promoter analysis, regulatory mechanisms, transcription factors.

Address Correspondence to:
Avi Ma'ayan
Department of Pharmacology
and Systems Therapeutics
Mount Sinai School of Medicine
New York, NY
Email: avi.maayan@mssm.edu

As progress in biomedical sciences leads to breakthrough in biotechnology, biotechnology is fueling back progress in biomedical sciences. However, as molecular data collection in mammalian cell biology research becomes more extensive and complex, advances in computation are now critical for continual progress.[1] Computational systems biology is playing an increasingly important role in many

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20192

© 2010 Mount Sinai School of Medicine

aspects of biomedical research.[2] Emerged as the result of the sequencing of the yeast, mouse, and human genomes, computational systems biology and bioinformatics (CSB) aims to develop robust theoretical models that explain how molecular components give rise to cellular, tissue, and organism phenotypes at the genome-wide scale.[3] Biotechnological advances in instruments capable of measuring molecular components within cells at a genome-wide level, together with the infusion of ideas from physics and mathematics for data analysis and advances from computer science for data sharing, storage, search, and visualization, are expected to lead to a surge in biomedical breakthroughs in translational research in the near future.

Biotechnological advances in instruments capable of measuring molecular components within cells at a genome-wide level, together with the infusion of ideas from physics and mathematics for data analysis and advances from computer science for data-sharing, storage, search, and visualization, are expected to lead to a surge in biomedical breakthroughs in translational research in the near future.

Although genome-wide proteomic approaches improve rapidly, currently and throughout the past decade the most widely available and cost-effective genome-wide expression data collected is at the RNA level. Typical studies examine cells under different experimental conditions such as control versus treated, or disease versus normal intracellular states. Since quantitative changes in mRNA levels do not directly explain how regulatory molecular mechanisms are altered to induce changes in gene expression, and in turn lead to changes in cellular phenotype, identifying such regulatory mechanisms has been the focus of many CSB studies. This is because such understanding will enable us to, among other things, better control cell behavior with small molecules, and in turn translate such ability to therapeutics development. Most popular approaches for data interpretation of changes in gene expression include promoter analysis,[4,5] gene ontology,[6] and pathway enrichment analyses,[7] as well as reverse engineering of networks from mRNA expression data.[8] The ultimate goal of many of these studies is to identify and rank potential target genes/proteins that if knocked down would explain the observed changes by reversing them. Such protein targets may ultimately become drug targets. In this article, we describe how protein kinases that are likely responsible for observed genome-wide changes in gene expression at the messenger RNA (mRNA) level can be identified and ranked by a rational approach. In the first phase, one combines promoter analysis of microarray results with data from various chromatin immunoprecipitation (ChIP-X) studies, such as chromatin immunoprecipitation sequencing (ChIP-Seq), chromatin immunoprecipitation coupled with paired-end ditag (ChIP-PET), and chromatin immunoprecipitation-on-chip (ChIP-chip). Then, using protein-protein interactions and kinase-protein phosphorylation reactions collected from the public domain, one can potentially identify and rank candidate protein kinases for knock-down and/or other types of functional validations.

In this article, we describe how protein kinases that are likely responsible for observed genome-wide changes in gene expression at the mRNA level can be identified and ranked by a rational approach. In the first phase, one combines promoter analysis of microarray results with data from ChIP-X studies. Then, using protein-protein interactions and kinase-protein phosphorylation reactions collected from the public domain, one can potentially identify and rank candidate protein kinases for knock-down and/or other types of functional validations.

We discuss how predictions made by this approach can be cross-validated using other types of high-throughput experimental data, such as phosphoproteomics and protein/DNA arrays, as well as using lists of protein kinases extracted automatically from the literature.

The overall workflow is shown schematically in Figure 1 and discussed below.

DOI:10.1002/MSJ

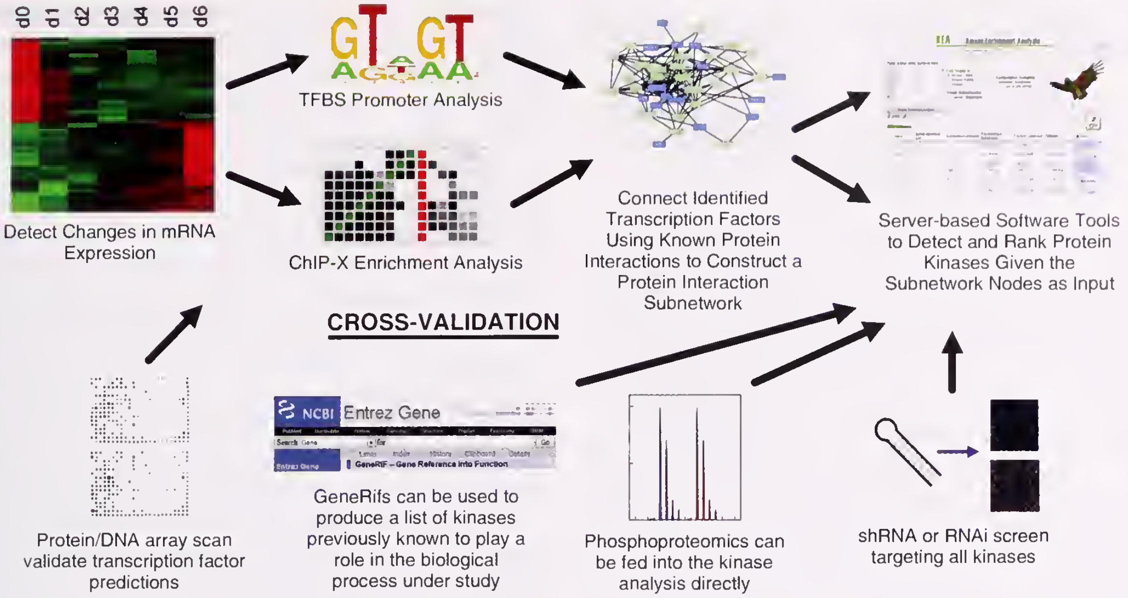

d0
d1
d2
d3
d4
d5
d6
Detect Changes in mRNA Expression
TFBS Promoter Analysis
ChIP-X Enrichment Analysis
Connect Identified Transcription Factors Using Known Protein Interactions to Construct a Protein Interaction Subnetwork
Server-based Software Tools to Detect and Rank Protein Kinases Given the Subnetwork Nodes as Input
CROSS-VALIDATION
Protein/DNA array scan validate transcription factor predictions
NCBI Entrez Gene
GeneRIF – Gene Reference into Function
GeneRifs can be used to produce a list of kinases previously known to play a role in the biological process under study
Phosphoproteomics can be fed into the kinase analysis directly
shRNA or RNAi screen targeting all kinases

Fig 1. Workflow process for inferring protein kinases from mRNA gene expression changes with added methods for cross-validation. **Abbreviations:** ChIP-X, chromatin immunoprecipitation studies; mRNA, messenger RNA; TFBS, transcription factor binding site.

DETECTING CHANGES IN MESSENGER RNA EXPRESSION

The first step of our CSB approach is to identify the genes (mRNAs) that are differentially expressed under 2 conditions. This is a standard procedure that can be carried out using statistical tests such as analysis of variance or an adjusted *t* test, and/or unsupervised clustering approaches such as hierarchical clustering or principal component analyses.

PROMOTER ANALYSIS AND CHROMATIN IMMUNOPRECIPITATION ENRICHMENT ANALYSIS

To link changes in gene expression to the molecular mechanisms responsible for the observed changes, we can first apply promoter analysis using binding site matrices obtained from databases such as TRANSFAC[4] or JASPAR.[5] This method computationally scans the DNA sequence in the proximity of genes' coding regions, seeking enrichment of binding sites for annotated transcription factor binding logo-motifs. Such an approach can identify and rank a list of transcription factor candidates responsible for the observed changes by computing binding-site enrichment for all the genes that changed in expression significantly. Transcription factor binding site enrichment can be computed for all genes that significantly changed in mRNA expression, or by dividing regulation for genes that were differentially increased or decreased in expression compared with the control. Alternatively, we can generate a list of most likely transcriptional regulators by cross-referencing the genes that increased or decreased in expression with previously published ChIP-X studies. Such studies report the binding of specific transcription factors in proximity to gene coding regions. By compiling the results from many ChIP-X studies, we can obtain a global picture of transcriptional activity of many transcription factors. Although such data is collected in many cell types and across different mammalian organisms under different conditions, it has the advantage that it considers the chromatin state of the cell and as such is expected to reduce false positives, a critical limitation of the binding logo-motif promoter scanning approach. Both the promoter analyses and the ChIP-X enrichment analyses produce ranked lists of transcription factors that most likely regulate genes that significantly increased or decreased in mRNA expression.

Both the promoter analyses and the ChIP-X enrichment analyses produce ranked lists of transcription factors that most likely regulate genes that significantly increased or decreased in mRNA expression.

Such lists can be compared for overlap to assess consistency.

CONNECTING IDENTIFIED TRANSCRIPTION FACTORS

Most analyses stop at this stage; however, our approach next step is to "connect" the transcription factors detected by the ChIP-X enrichment and/or by the promoter scanning approach using known experimentally reported protein-protein interactions. Several tools have been developed for using prior knowledge about protein-protein interaction networks to build subnetworks that connect lists of "seed nodes" given as input.[9–12] We have developed Genes2Networks[10] and successfully used it for finding pathways responsible for neurite outgrowth[13] and predicting a novel disease gene, SHOC2, that was found to contain a mutation that can cause a Noonan-like syndrome.[14] Once we have built a subnetwork that connects the transcription factors, we can convert the proteins (nodes) within this subnetwork to a list (of proteins). Such a list of proteins can then be fed into Kinase Enrichment Analysis (KEA), a Web tool we developed for identifying protein kinases that are enriched with substrates in a given input list of proteins[15]; or into sequence-based motif discovery tools such as Motif-X[16] or GPS[17] for identifying the protein kinases that most likely regulate the subnetwork anchored with the transcription factors detected to regulate the observed changes in gene expression. Alternatively, programs such as Motif-X or similar sequence-based approaches[17,18] can be used to associate protein kinases enriched with substrate sequences in the subnetwork. The final output of the workflow is a ranked list of protein kinases as targets for experimental validation through knock-down, overexpression, or dominant negative perturbations.

The final output of the workflow is a ranked list of protein kinases as targets for experimental validation through knock-down, overexpression, or dominant negative perturbations.

VARIATIONS IN WORKFLOW

Some of the steps in the process can be skipped. We can, for example, look for kinases that significantly changed in their mRNA level directly. Alternatively, we can look for kinases in the protein-protein interactions subnetworks created from the transcription factors detected computationally, avoiding the KEA or Motif-X analyses. We also have the option to only use transcription factor binding motif promoter scanning and not ChIP-X, or to use only KEA and not Motif-X. How can we evaluate which approach works best? To answer this question we can bring in additional experimental evidence for cross-validation. For example, if we have data describing changes in protein/DNA interactions,[13] or proteomics data reporting quantitative changes in protein levels (eg, profiling the entire nuclear proteome[19]), or, more directly, stable isotope labeling with amino acids in cell culture (SILAC) phosphoproteomics data describing quantitative changes in phosphoprotein levels,[20] we can evaluate and validate the different steps of the analysis. Protein/DNA interaction arrays such as those developed by Panomics/Affymetrix (www.panomics.com/index.php?id=product_18) can validate the transcription factors detected by the promoter analysis or the ChIP-X enrichment analysis. SILAC phosphoproteomics followed by a KEA or Motif-X analysis can validate the final output of the protein kinase rankings. In addition, we can see if the candidate protein kinases detected through the analysis were previously implicated with the biological process we aim to better understand. Ultimately, we can employ an RNA interference (RNAi) screen targeting all kinases to see if our rankings reflect our predictions. The different validation approaches are summarized in the schematic provided in Figure 1 under the title Cross-Validation. All the computational analyses presented here can be achieved by meshing the programs Genes2Networks,[10] Lists2Networks,[21] KEA,[15] and ChEA (unpublished) developed by the Ma'ayan laboratory.

All the computational analyses presented here can be achieved by meshing the programs Genes2Networks, Lists2Networks, KEA, and ChEA (unpublished) developed by the Ma'ayan laboratory.

CONCLUSION

Protein kinases are great candidate drug targets because they display global effects through their ability to phosphorylate many targets exhibiting broad changes in cell behavior. Small-molecule inhibitors that target protein kinases can be easily screened for hits through in vitro phosphorylation assays. Protein kinases contain regulatory structural domains as well as catalytic domains amenable for small-molecule targeting. Here we have described several alternatives for inferring and ranking protein kinases, both computationally and experimentally, and for cross-validation of predictions initially made based on gene expression changes. However, there is still a need to validate the approach presented experimentally. With our approach there are many "free parameters" as well as "process routes" for making the ranked kinase predictions. Hence, protein kinase rankings can result in many different output tables depending on the computational and/or experimental routes taken and parameters chosen for cutoffs and algorithmic setup in each step. Although cross-validation could be useful for directing the right route and parameter selection, experimental validation, through functional studies as well as through short hairpin RNA (shRNA) or RNAi screens, is probably the best approach for evaluating the approach presented. CSB is now at a stage where data integration across regulatory layers can be used to infer and unravel regulation at the molecular level of mammalian cells at a degree of detail never possible before. Computational approaches for data integration and creative algorithms that can extract new knowledge from the increasingly complex, multidimensional, diverse datasets are of critical importance in this milieu.

ACKNOWLEDGEMENTS

The research presented is supported by NIH grants R01-DK088541 to JCH and AM, R01-DK078897 to JCH, and P50-GM071558-01A27398 to AM.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Ma'ayan A. Insights into the organization of biochemical regulatory networks using graph theory analyses. *J Biol Chem* 2009; 284: 5451–5455.
2. Ma'ayan A. Network integration and graph analysis in mammalian molecular systems biology. *IET Syst Biol* 2008; 2: 206–221.
3. Ma'ayan A, Blitzer RD, Iyengar R. Toward predictive models of mammalian cells. *Annu Rev Biophys Biomol Struct* 2005; 34: 319–349.
4. Matys V, Kel-Margoulis OV, Fricke E, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res* 2006; 34(suppl 1): D108–D110.
5. Portales-Casamar E, Thongjuea S, Kwon AT, et al. JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2010; 38(suppl 1): D105–D110.
6. The Gene Ontology Consortium. The Gene Ontology in 2010: extensions and refinements. *Nucleic Acids Res* 2010; 38(suppl 1): D331–D335.
7. Kanehisa M, Goto S, Furumichi M, et al. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 2010; 38(suppl 1): D355–D360.
8. Margolin A, Nemenman I, Basso K, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 2006; 7(suppl 1): S7.
9. Wu J, Vallenius T, Ovaska K, et al. Integrated network analysis platform for protein-protein interactions. *Nat Methods* 2009; 6: 75–77.
10. Berger SI, Posner JM, Ma'ayan A. Genes2Networks: connecting lists of gene symbols using mammalian protein interactions databases. *BMC Bioinformatics* 2007; 8: 372.
11. Lee SA, Chan CH, Chen TC, et al. POINeT: protein interactome with sub-network analysis and hub prioritization. *BMC Bioinformatics* 2009; 10: 114.
12. Antonov AV, Dietmann S, Rodchenkov I, Mewes HW. PPI spider: a tool for the interpretation of proteomics data in the context of protein-protein interaction networks. *Proteomics* 2009; 9: 2740–2749.
13. Bromberg KD, Ma'ayan A, Neves SR, Iyengar R. Design logic of a cannabinoid receptor signaling network that triggers neurite outgrowth. *Science* 2008; 320: 903–909.
14. Cordeddu V, Di Schiavi E, Pennacchio LA, et al. Mutation of SHOC2 promotes aberrant protein N-myristoylation and causes Noonan-like syndrome with loose anagen hair. *Nat Genet* 2009; 41: 1022–1026.
15. Lachmann A, Ma'ayan A. KEA: kinase enrichment analysis. *Bioinformatics* 2009; 25: 684–686.
16. Schwartz D, Gygi SP. An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat Biotechnol* 2005; 23: 1391–1398.
17. Xue Y, Zhou F, Zhu M, et al. GPS: a comprehensive www server for phosphorylation sites prediction. *Nucleic Acids Res* 2005; 33(suppl 2): W184–W187.
18. Linding R, Jensen LJ, Pasculescu A, et al. NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res* 2008; 36(suppl 1): D695–D699.
19. Lu R, Markowetz F, Unwin RD, et al. Systems-level dynamic analyses of fate change in murine embryonic stem cells. *Nature* 2009; 462: 358–362.
20. Blagoev B, Ong SE, Kratchmarova I, Mann M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat Biotechnol* 2004; 22: 1139–1145.
21. Lachmann A, Ma'ayan A. Lists2Networks: integrated analysis of gene/protein lists. *BMC Bioinformatics* 2010; 11: 87.

Mount Sinai Journal of Medicine 77:350–357, 2010

Integration of Small-Molecule Discovery in Academic Biomedical Research

Michael Ohlmeyer, PhD, and Ming-Ming Zhou, PhD

Department of Structural and Chemical Biology, Mount Sinai School of Medicine, New York, NY

OUTLINE

CHALLENGES AND OPPORTUNITIES FOR TARGETING CELLULAR PROTEINS
CHEMICAL SPACE FOR SMALL-MOLECULE DISCOVERY
HIT-FINDING AND HIGH-THROUGHPUT SCREENING
HIT-TO-LEAD AND MEDICINAL CHEMISTRY
  Hit Confirmation
  Hit-to-Lead Chemistry
  Lead Optimization
CONCLUSION
SUPPORTING INFORMATION

ABSTRACT

Rapid advances in biomedical sciences in recent years have drastically accelerated the discovery of the molecular basis of human diseases. The great challenge is how to translate the newly acquired knowledge into new medicine for disease prevention and treatment. Drug discovery is a long and expensive process, and the pharmaceutical industry has not been very successful at it, despite its enormous resources and spending on the process. It is increasingly realized that academic biomedical research institutions ought to be engaged in early-stage drug discovery, especially when it can be coupled to their basic research. To leverage the productivity of new-drug development, a substantial acceleration in validation of new therapeutic targets is required, which would require small molecules that can precisely control target functions in complex biological systems in a temporal and dose-dependent manner. In this review, we describe a process of integration of small-molecule discovery and chemistry in academic biomedical research that will ideally bring together the elements of innovative approaches to new molecular targets, existing basic and clinical research, screening infrastructure, and synthetic and medicinal chemistry to follow up on small-molecule hits. Such integration of multidisciplinary resources and expertise will enable academic investigators to discover novel small molecules that are expected to facilitate their efforts in both mechanistic research and new-drug target validation. More broadly academic drug discovery should contribute new entities to therapy for intractable human diseases, especially for orphan diseases, and hopefully stimulate and synergize with the commercial sector. ***Mt Sinai J Med 77:350–357, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: chemical screening, drug discovery, medicinal chemistry, small molecule.

Address Correspondence to:

Ming-Ming Zhou
Department of Structural and
Chemical Biology
Mount Sinai School of Medicine
New York, NY
Email: ming-ming.zhou@mssm.edu

The completion of the human genome sequencing project has fundamentally changed the way scientific research is conducted today. The genetic information of the human genome holds the key to our most fundamental questions in human biology of health and disease. However, despite the rapidly accumulating knowledge defining the molecular basis of human disease and the surge in the number of new potential therapeutic targets, annual new-drug approvals by the US Food and Drug Administration (FDA) have been steadily decreasing in the last decade, to <20 new drugs in 2008–2009.

It is well known that drug discovery is a long and very expensive process (on average, ≥15 years and >$1 billion) and that the pharmaceutical industry has not been always successful at it, despite its enormous resources and spending on the process. Many failures occur in the later stages of the drug-discovery process that are due to lack of efficacy or intolerable toxicity in humans, making such failures very costly. Therefore, to leverage the productivity of new-drug development, a substantial acceleration in validation of potential therapeutic targets is required. Target

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20197

© 2010 Mount Sinai School of Medicine

validation, however, requires tremendous efforts in basic and mechanism-driven research at the molecular and cellular levels, as well as in animal models. It has become increasingly realized that academic medical research institutions ought to be engaged in early-stage drug discovery, especially when it can be coupled to their basic research activities[1–3] Ideally, validation of potential new-drug targets would also benefit enormously from sustained interactions between academic laboratories and the pharmaceutical industry. Unfortunately, because of the high-risk nature and the extent of the basic mechanistic research involved, the pharmaceutical companies are generally not enthusiastic about early-stage target validation, nor does a typical academic institution have access to high-throughout screening (HTS) capabilities to routinely obtain commercially viable lead compounds for proof-of-principle target validation. Consequently, some major biomedical research institutions in the United States in recent years have started developing academic drug-discovery programs. These programs aim to facilitate the identification of small-molecule chemical compounds to be used for validation of potential new therapeutic targets being discovered in basic research. In addition, these small-molecule agents, which are capable of modulating the biological activity of a disease-relevant protein, enzyme, or macromolecule, can also serve as powerful research tools that are complementary to other technologies, such as site-directed mutagenesis and gene knockdown by RNA interference.

To leverage the productivity of new-drug development, a substantial acceleration in validation of potential therapeutic targets is required. For this reason, academic medical research institutions ought to be engaged in early-stage drug discovery, especially when it can be coupled to their basic research activities.

The advantages of perturbing biological functions of macromolecular protein targets by using small molecules as chemical probes as compared with the other techniques result from the ability of small molecules to modulate selectively individual functions of a target protein in its endogenous form under physiological or pathophysiological conditions. Moreover, effects of such small-molecule probes can often be controlled precisely in a temporal and dose-dependent manner, which makes them ideal research tools for investigating biological mechanisms, and certainly for validating potential new-drug targets in complex biological systems. Therefore, small-molecule discovery should not only facilitate early-stage research in drug discovery through new target validation, but also advance mechanistic understanding of human biology of health and disease. Integration of small-molecule discovery and chemistry is increasingly possible in academic research, and here we will outline some of the opportunities and challenges in this process.

The effects of small-molecule probes can often be controlled precisely in a temporal and dose-dependent manner, which makes them ideal research tools for investigating biological mechanisms, and certainly for validating potential new-drug targets in complex biological systems.

CHALLENGES AND OPPORTUNITIES FOR TARGETING CELLULAR PROTEINS

Given the rapid advances in our understanding of structure-function of protein families in the human genome in general and the available technologies to gain specific knowledge for a given protein target, it has become apparent that druggable[4,5] biological targets could include those beyond the usual classes, such as enzymes and membrane-bound receptors, that have been the traditional focus of the pharmaceutical industry. New target classes include protein modular domains that engage in protein-protein interactions,[6] which are modulated by posttranslational modifications of amino acids. Examples are the bromodomains and chromodomains, which are present in chromatin-modifying enzymes and associated proteins and function to recognize acetylated and methylated lysine residues, respectively, in DNA-packing histones and transcriptional proteins during gene transcription[7,8] Small-molecule targeting of such protein interaction domains has advantages as compared with that of enzymes, because these protein domains, which serve as functional units within large multidomain proteins, exert their activity by interacting with their binding partners with high specificity but not necessarily high affinity individually (K_d of 1–100 μM)[9–11] This unique feature of these protein domains allows rapidly reversible and dynamic regulation of cellular biology, such as signal transduction and gene transcription[12] Therefore, targeting of amino acid modification–mediated protein interactions with selective, high-affinity small molecules provides a new mechanism for modulation of biological activities and pathways.[13–15]

> *Targeting of amino acid modification–mediated protein interactions with selective, high-affinity small molecules provides a new mechanism for modulation of biological activities and pathways.*

Because of the dynamic nature of protein-protein interactions, in that target protein–biological ligand binding interface is not preformed in the absence of the ligand, effective lead discovery would require a screening assay, in which the target protein exists in a state mimicking the ligand-bound state. As such, target structure-guided screening is a more effective and rational approach to ligand design, as it takes advantage of the three-dimensional structure of the target, including the structure of the crucial functional site(s), which can guide one to build specific chemical entities that interact with these sites[16–19] The novel ligands can be built with chemical entities constructed from building blocks identified from target structure-based screening and optimization coupled with focused library synthesis. The resulting linked compounds with high affinity and selectivity are then subject to detailed structure-activity relationship analysis of their interactions with the target by using nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography, mass spectrometry, and computational methods. Refinement, chemical diversification through chemical linkages, and selectivity enhancement can be further achieved at this stage.

CHEMICAL SPACE FOR SMALL-MOLECULE DISCOVERY

It has been estimated that theoretically as many as 10^{60} different chemical molecules could be in existence,[20] highlighting a vastness of chemical space that is matched only by that of the cosmological universe. However, due to the limits of current synthetic chemistry technologies, not all theoretic chemical species could be made, nor are they all necessary for identifying small molecules for a given biological target. In fact, studies have suggested that biologically active compounds, which have favorable physicochemical properties for interactions with biological molecules such as proteins and nucleic acids, are sparsely clustered in the chemical universe according to the target families. Therefore, it is thought that an ideal chemical library to be used for chemical screening would consist of small molecules that represent these discrete regions of biologically relevant chemical space. Chemical compounds could then be selected from a large collection of available chemical compounds using molecular descriptors such as pharmacophores, which describe key spatial arrangement and physicochemical properties related to their potential biological activity. Therefore, an ideal chemical library to be used for HTS should consist of small molecules with chemical diversity covering biological target families as well as drug-like physicochemical properties. Chemical hits and lead compounds identified from chemical library HTS can then serve as starting points for target-based medicinal chemistry that will develop these lead compounds into potent and selective chemical probes for a given target (see Figure 1).

Integration of discovery and biological applications of small-molecule probes can enhance the

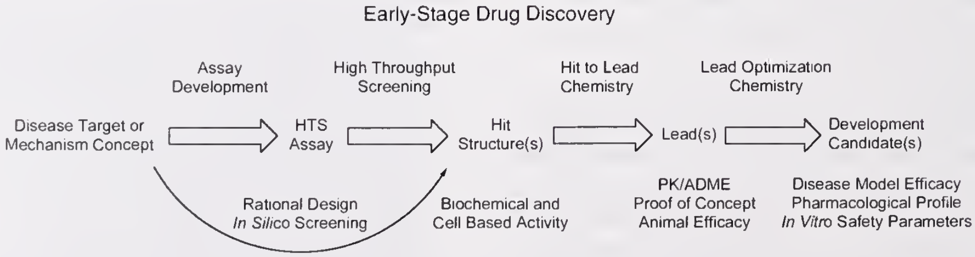

Early-Stage Drug Discovery
Assay Development
High Throughput Screening
Hit to Lead Chemistry
Lead Optimization Chemistry
Disease Target or Mechanism Concept
HTS Assay
Hit Structure(s)
Lead(s)
Development Candidate(s)
Rational Design In Silico Screening
Biochemical and Cell Based Activity
PK/ADME Proof of Concept Animal Efficacy
Disease Model Efficacy Pharmacological Profile In Vitro Safety Parameters

Fig 1. Schematic illustration of key steps involved in early-stage small-molecule drug discovery.

development of biologically active compounds. It can also enable researchers to effectively investigate the underlying molecular mechanisms of important cellular processes in human biology and disease, and as well help discover and validate new therapeutic targets for disease treatment[20,21] Given that the academic small-molecule discovery is often aimed to achieve objectives of both aiding mechanistic research and validation of new therapeutic targets, the process can be different from the traditional linear drug-discovery process. In the latter case, potential new-drug targets are known based on the knowledge of a given disease condition. Compounds in a chemical library are evaluated in a high-throughput assay for their ability to interact with and functionally modulate the protein target. Initial chemical hits, which exhibit activity above a defined threshold, are further optimized for their activity through chemical modifications to yield chemical leads that possess the required pharmacokinetic properties. These leads are subsequently evaluated in in vivo biological and disease models and undergo lead optimization through additional chemical modifications. The latter process is performed in an iterative fashion, developing the lead into clinical drug candidates that are finally evaluated in human clinical trials before being subjecting to FDA review and approval as new therapeutic agents for disease treatment.

> *Integration of discovery and biological applications of small-molecule probes can enhance the development of biologically active compounds.*

Given that basic research of interest may or may not be linked directly to a particular disease model, chemical screening can still be conducted to identify small molecules that elicit specific phenotypic effects, such as cell-cycle arrest or apoptosis, or modulate in vitro functions of a target protein of interest. Such chemical agents can be used as research probes in chemical genetics research to identify biological targets responsible for the phenotypic effects, or validate specific proteins as potential therapeutic targets for disease treatment. Alternatively, when the chemical screening is conducted using different disease models, emerging small-molecule chemical agents could be used to dissect the underlying molecular mechanisms of the disease models and also lay the foundation for further development of new therapeutic agents.

HIT-FINDING AND HIGH-THROUGHPUT SCREENING

Starting compounds for the small-molecule drug discovery process are conventionally termed "hits"; these are compounds with defined, reproducible activity in a primary, often biochemical (eg, receptor binding, enzyme inhibition), screening assay (Figure 1). There are several approaches to hit-finding,[22–24] including de novo structure-based design, use of peptides or other endogenous ligands, and screening approaches. These are not mutually exclusive and may be pursued in parallel depending on the maturity of a project or target area. Screening approaches are of interest because of their general applicability to all types of targets at a very early stage. Since at least the early 1990s, the use of HTS has been a standard component in the hit-finding process in the pharmaceutical industry, where "high throughput" conventionally refers to testing of several tens of thousands (or more) of compounds in the course of a few weeks or months.

Two elements are required for HTS: (1) a compound collection to screen, and (2) assays suitable for high throughput and the infrastructure and/or hardware to implement them. In recent years both of these elements have become accessible to academic research institutions. Partly in response to the demands of the pharmaceutical industry (and the economic possibilities afforded by outsourcing), it has become possible for academic institutions to assemble significant screening collections (>100, 000 compounds) from commercial sources (eg, ChemBridge, ChemDiv, and Maybridge). The Mount Sinai collection currently comprises approximately 115,000 compounds prescreened for physicochemical properties appropriate for drug discovery,[21,25] for example Lipinski's Rule of Five.[26] The Rule of Five[26] was established based on an analysis of physicochemical properties of drug molecules that reached phase II clinical trials. The rule states that the drug-like properties of a chemical compound should consist of (1) <5 hydrogen bond donors (sum of OHs and NHs), (2) molecular weight <500 Da, (3) a Log P (partition coefficient, a measure of hydrophobicity) <5, (4) <10 hydrogen bond acceptors (sum of N and O atoms), and (5) compound classes that are substrates for biological transporters are an exception to the rule[26]

Several further points are worth noting. First, a screening collection is not a static entity, but may evolve in several ways, including addition (either by purchase or in-house synthesis) of focused sets of compounds directed at particular target classes, for example kinases, nuclear hormone receptors, or

G-protein-coupled receptors. Libraries built around pharmacologically privileged scaffolds[27] and natural product–like[28] libraries may be added. Second, the collection will evolve naturally as compounds are added from Mount Sinai hit-to-lead and lead-optimization programs, and this provides significant opportunities because proprietary structural motifs exploited versus one target may be of use to discover hits in an unrelated area. One particular opportunity afforded by an in-house screening collection coupled to active, ongoing research programs is the use of in silico docking methods to prioritize and focus screening (where there is sufficient structural information about targets).[29]

The second element in mounting an HTS campaign are assays suitable for high throughput and the infrastructure necessary to run them. In vitro assays emerging from laboratories doing basic biological research often involve multiple reagent additions, transfer steps, washes, or filtrations. This type of protocol, though acceptable for daily or weekly evaluation of a handful of compounds, is not practical for HTS, where on the order of hundreds of thousands of assay points will be required to complete the survey of even a modest compound collection. The optimal assay strategy for HTS is a "mix-and-read" format in which biological components can be combined in microtiter plates, reagents added in 1 or 2 steps, and then a signal read from the assay plate directly without filtrations, washes, or separations. Again, partly in response to the HTS demands of the pharmaceutical industry, a wide variety of assay types is now possible in a mix-and-read format, including receptor binding assays, protein-protein binding, enzyme assays, and cell-signaling assays.[30] Readouts are conventionally based on fluorescence, fluorescence polarization, bioluminescence, or chemiluminescence, with a variety of commercial sources for reagent. The robotics systems, liquid handling, and plate readers for conventional HTS are all within the reach of major academic research institutions.

In addition to in-house compound collections and screening facilities, academic researchers may also access the Molecular Library Screening Centers Network (MLSCN, part of the National Institutes of Health [NIH] Roadmap initiative), in which an assay is transferred to the MLSCN and screened versus a library of approximately 150,000 diverse compounds. Assay results are then deposited in PubChem. Entry for assays to MLSCN is via a peer-review process.[31]

Beyond screening for single biomolecular interactions or transformations, an emerging technique is the use of cellular imaging in high-throughput mode in what has been termed high-content screening.[32,33] Here, whole cells are imaged in tandem with screening-compound challenge to allow simultaneous readout of multiple parameters, for example binding, downstream events relating to mechanism of action (eg, translocation of transcription factors), and markers of cell viability, which provide an early indication of potential toxicity.

HIT-TO-LEAD AND MEDICINAL CHEMISTRY

The result of an HTS campaign is typically a small group of structures that show activity in the primary screening assay. At this stage, synthetic and medicinal chemists can begin to interact with their fellow scientists in biology and pharmacology in the interactive process of taking a screening hit to a medicinal chemistry lead, and further to a clinical candidate. Academic centers committed to exploiting chemical biology to develop tools for basic research[34–36] and also aim to engage in drug discovery have set up chemistry groups dedicated to medicinal chemistry, either alongside their traditional academic chemistry departments or de novo, as at Mount Sinai. As mentioned earlier, the drug-discovery process is a collaborative enterprise requiring input from a range of scientific disciplines–indeed, this is one of the organizational challenges for drug discovery in an academic environment, which traditionally has been structured around research groups directed by individual investigators. Next we outline the early part of the compound-optimization process from the perspective of progression of a compound from screening hit to medicinal chemistry lead–from a chemist's perspective.

Hit Confirmation

Data from an HTS campaign are typically triaged to eliminate false positives or artifacts that might arise from, for example, interfering-compound fluorescence in a fluorescence polarization assay[37] or compound protein-aggregation artifacts.[38,39] Compound quantities in screening collections are generally small and often will be available only as a dimethyl sulfoxide stock solution. Thus, putative active structures are typically resynthesized and scaled up to confirm identity and purity. Hit compounds can then be made available to biologists in adequate quantity to be interrogated in detail in the primary screening assay and other follow-up assays, often cell based, to confirm useful levels of activity. This is the first, simplest, but often crucial step in follow-up of an HTS campaign, or it might serve as the first experimental confirmation of an in silico (rational) design approach to hit-finding.

Hit-to-Lead Chemistry

Confirmed hits are assessed and ranked based on a number of factors in addition to potency, particularly chemical tractability. A chemically tractable lead will be of reasonably low molecular-weight structure (<500 Da), which is amenable to rapid analog synthesis to facilitate exploration of structure-activity relationships. Depending on the molecular type, analogs may be available for purchase from commercial sources; however, generally, parallel libraries will be designed and synthesized to explore various sites for modification and improvement of the structure. This process may go through several cycles of parallel library synthesis and testing, with each iteration focusing on and optimizing a feature or site on the hit structure. Several relatively new tools are available to medicinal chemists to facilitate rapid analog generation, for example microwave-assisted synthesis,[40] which allows reaction times to be shortened (and often gives fewer unwanted byproducts) and polymer-assisted solution phase synthesis, either as reagents or scavengers, which simplifies reaction workup and facilitates purification.[41,42] All or most of the molecules from a hit-to-lead synthesis are likely to be suitable for inclusion in the broader screening library and so contribute to its growth and evolution.

A successful screening campaign will ideally yield 3 or 4 tractable lead series suitable for this type of early exploration. Compound series where potency and selectivity can be usefully modulated will be characterized in more detail with respect to physicochemical parameters such as solubility log P to confirm theoretical or calculated properties and confirm drug-like characteristics. In major academic research institutions it may be possible to form collaborations with established research groups, for example in computational chemistry or structural biology to test theoretical models or aid in molecular design. This is the case at Mount Sinai, where there is established expertise in computational biology, molecular simulation, and experimental structure determination by using NMR and X-ray crystallography. A second institutional and organizational benefit is that medicinal chemistry might be applied to scientifically important projects in an academic setting, whereas it may not be possible in the pharmaceutical industry for purely business reasons (potential market size). Advanced molecules from tractable hit series where in vitro potency (in biochemical and cell-based assays) and selectivity can be improved serve as valuable research tools for interrogation of novel disease targets and may help validate novel or speculative biological mechanisms as potential drug targets–either of which may be a scientifically reasonable and successful endpoint for academic research (which is emphatically not the case in commercial drug discovery). Advanced structures from hit-to-lead chemistry may also form the jumping-off point for further optimization.

Lead Optimization

Based on an assessment of biochemical potency, cell-based activity, and overall physicochemical profile, one or more series of structures may be selected for further optimization. The objectives of lead optimization are to increase potency in cell-based assays and to test molecules in more downstream in vitro and in vivo functional assays. In vitro absorption, distribution, metabolism and excretion (ADME) parameters such as microsomal stability and physical properties such as aqueous solubility will be monitored as leads progress.[43] The goal is an adequate balance of potency and physicochemical properties such that early rodent in vivo exposure data (pharmacokinetics) can be obtained. Early pharmacokinetics assessment has two goals: (1) to confirm that the series does indeed have potential for acceptable bioavailability, and (2) to provide compound exposure data to help in the design and interpretation of initial in vivo efficacy experiments. In an academic setting where drug-discovery targets are likely to be natural outgrowths and extensions of established basic research programs, access to in vivo efficacy models, or in-house expertise in setting them, is likely to be an advantage. At the stage in lead optimization where promising proof-of-concept in vivo efficacy data in animal models is obtained, broader pharmacological profiling of leads is appropriate, for example, selectivity profiling versus panels of receptors and enzymes. This type of broad pharmacological profile is available from contract service organizations such as Cerep and Caliper Life Sciences. Early in vitro safety parameters may also be assessed, for example hERG channel activity (an indicator of potential cardiovascular liabilities) or cytochrome P450 activity (an indicator of potential drug-dug interactions).[44] This type of profiling can be reliably outsourced to contract research organizations that service the pharmaceutical and biotech industries. Academic drug-discovery projects may utilize the same type of resources, or, for NIH-funded projects, resources such as the Molecular Pharmacology Research Program of the National Institute of Mental Health. Lead profiling of this type will identify potentially problematic off-target activities, which can be addressed with continued optimization to either ameliorate or remove them (or result in a no-go decision on a lead series). Optimized leads with in vivo efficacy in animal models (either disease or biomarker), at an acceptable dose and route of administration, with no obvious toxicity or metabolic liabilities, qualify as candidates for more long-term and detailed (including regulatory) animal toxicology studies and are sometimes designated as development candidates and conventionally regarded as the output of early drug discovery.

CONCLUSION

Integration of small-molecule discovery and chemistry in academic biomedical research will ideally bring together the elements of (1) innovative approaches to new molecular targets (such as protein domain interactions), (2) existing basic and clinical research, (3) screening infrastructure, and (4) synthetic and medicinal chemistry to follow up on small-molecule hits. Such integration of multidisciplinary resources and expertise will enable academic investigators to discover novel small molecules that are expected to facilitate their efforts in both mechanistic research and new-drug target validation. More broadly academic drug discovery should contribute new entities for therapy (especially in neglected and orphan disease areas) and potentially stimulate and synergize with the commercial sector.

SUPPORTING INFORMATION

The authors wish to acknowledge the support by Mount Sinai School of Medicine. The work was supported by grants from the National Institutes of Health (M.-M.Z.).

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Verkman AS. Drug discovery in academia. *Am J Physiol Cell Physiol* 2004; 286: C465–C474.
2. Wyatt PG. The emerging academic drug-discovery sector. *Future Med Chem* 2009; 1: 1013–1017.
3. Kozikowski AP, Roth B, Tropsha A. Why academic drug discovery makes sense. *Science* 2006; 313: 1235–1236.
4. Hopkins AL, Groom CR. The druggable genome. *Nat Rev Drug Discov* 2002; 1: 727–730.
5. Kubinyi H. Drug research: myths, hype and reality. *Nat Rev Drug Discov* 2003; 2: 665–668.
6. Berg T. Modulation of protein-protein interactions with small organic molecules. *Angew Chem Int Ed Engl* 2003; 42: 2462–2481.
7. Sanchez R, Zhou MM. The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Devel* 2009; 12: 659–665.
8. Taverna SD, Li H, Ruthenberg AJ, et al. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 2007; 14: 1025–1040.
9. Mujtaba S, He Y, Zeng L, et al. Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain. *Mol Cell* 2002; 9: 575–586.
10. Mujtaba S, He Y, Zeng L, et al. Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol Cell* 2004; 13: 251–263.
11. Fischle W, Wang Y, Jacobs SA, et al. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev* 2003; 17: 1870–1881.
12. Huang B, Yang XD, Zhou MM, et al. Brd4 coactivates transcriptional activation of NF-kappaB via specific binding to acetylated RelA. *Mol Cell Biol* 2009; 29: 1375–1387.
13. Zeng L, Li J, Muller M, et al. Selective small molecules blocking HIV-1 Tat and coactivator PCAF association. *J Am Chem Soc* 2005; 127: 2376–2377.
14. Sachchidanand, Resnick-Silverman L, Yan S. Target structure-based discovery of small molecules that block human p53 and CREB binding protein association. *Chem Biol* 2006; 13: 81–90.
15. Pan C, Mezei M, Mujtaba S, et al. Structure-guided optimization of small molecules inhibiting human immunodeficiency virus 1 Tat association with the human coactivator p300/CREB binding protein-associated factor. *J Med Chem* 2007; 50: 2285–2288.
16. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 1996; 274: 1531–1534.
17. Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins. *Science* 1997; 278: 497–499.
18. Hajduk PJ, Gerfin T, Boehlen JM, et al. High-throughput nuclear magnetic resonance-based screening. *J Med Chem* 1999; 42: 2315–2317.
19. Moore JM. NMR screening in drug discovery. *Curr Opin Biotechnol* 1999; 10: 54–58.
20. Dobson CM. Chemical space and biology. *Nature* 2004; 432: 824–828.
21. Lipinski C, Hopkins A. Navigating chemical space for biology and medicine. *Nature* 2004; 432: 855–861.
22. Goodnow RA Jr, Gillespie P. Hit and lead identification: efficient practices for drug discovery. *Prog Med Chem* 2007; 45: 1–61.
23. Keseru GM, Makara GM. Hit discovery and hit-to-lead approaches. *Drug Discov Today* 2006; 11: 741–748.

24. Gribbon P. High-throughput hit finding and compound-profiling technologies for academic drug discovery. *Drug Discov Today Tech* 2008; 5: e3–e7.
25. Leeson PD, Springthorpe B. The influence of drug-like concepts on desision-making in medicinal chemistry. *Nat Rev Drug Discov* 2007; 6: 881–890.
26. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 2001; 46: 3–26.
27. Horton DA, Bourne GT, Smythe ML. The combinatorial synthesis of bicyclic privileged structures or priveleged substructures. *Chem Rev* 2003; 103: 893–930.
28. Nandy JP, Prakesch M, Khadem S, et al. Advances in solution-and solid-phase synthesis toward the generation of natural product-like libraries. *Chem Rev* 2009; 109: 1999–2060.
29. Morris GM, Goodsell DS, Halliday RS, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 1998; 19: 1639–1662.
30. Inglese J, Johnson RL, Simeonov A, et al. High-throughput screening assays for the identification of chemical probes. *Nat Chem Biol* 2007; 3: 466–479.
31. Lazo JS, Brady LS, Dingledine R. Building a pharmacological lexicon: small molecule discovery in academia. *Mol Pharmacol* 2007; 72: 1–7.
32. Lang P, Yeow K, Nichols A, Scheer A. Cellular imaging in drug discovery. *Nat Rev Drug Discov* 2006; 5: 343–356.
33. Krausz E. High-content siRNA screening. *Mol Biosyst* 2007; 3: 232–240.
34. Stockwell BR. Exploring biology with small organic molecules. *Nature* 2004; 432: 846–854.
35. Bauer A, Stockwell B. Neurobiological applications of small molecule screening. *Chem Rev* 2008; 108: 1774–1786.
36. Spring DR. Chemical genetics to chemical genomics: small molecules offer big insights. *Chem Soc Rev* 2005; 34: 472–482.
37. Owicki JC. Fluoresence polarization and anisotropy in high throughput screening: perspectives and primer. *J Biomol Screen* 2000; 5: 297–306.
38. Shoichet BK. Screening in a spirit haunted world. *Drug Discov Today* 2006; 11: 607–615.
39. McGovern SL, Helfand BT, Feng B, Shoichet BK. A specific mechanism of nonspecific inhibition. *J Med Chem* 2003; 46: 4265–4272.
40. Raveglia LF, Giardina GAM. Accelerating the drug-discovery process: new tools and technologies available to medicinal chemists. *Future Med Chem* 2009; 1: 1019–1023.
41. Kirschning A, Monenschein H, Wittenberg R. Functionalized polymers–emerging versatile tools for solution-phase chemistry and automated parallel synthesis. *Angew Chem Int Ed Engl* 2001; 40: 650–679.
42. Ley SV, Baxendale IR. New tools and concepts for modern organic synthesis. *Nat Rev Drug Discov* 2002; 1: 573–586.
43. Wang J, Urban L, Bojanic D. Maximising use of in vitro ADMET tools to predict in vivo bioavailability and safety. *Expert Opin Drug Metab Toxicol* 2007; 3: 641.
44. Kramer JA, Sagartz JE, Morris DL. The application of discovery toxicology and pathology towards the design of safer pharmaceutical lead candidates. *Nat Rev Drug Discov* 2007; 6: 636–649.

Pursuit of Personalized Anticancer Therapy: Leveraging Collaboration Between Academia and the Biotech/Pharmaceutical Industry

Elizabeth Buck, PhD, Mark Mulvihill, PhD, and Kenneth K. Iwata, PhD

OSI Pharmaceuticals, Farmingdale, NY

OUTLINE

Development of Cetuximab
Development of Herceptin
Targeting BCR-ABL for the Treatment of Chronic Myelogenous Leukemia
Discovery of Osi-906
Conclusion

ABSTRACT

Over the past 2 decades, our increased understanding of tumor biology has resulted in the delivery of a new generation of molecularly targeted cancer drugs with greater efficacy and less toxicity. This understanding has also provided pharmaceutical and academic institutions with a greater appreciation for the complexities and challenges associated with discovering and developing molecularly targeted drugs. To deal with the complexities of tumor biology and the associated technologies needed to develop molecularly targeted drugs, there has been increased cooperation and collaboration between academic and pharmaceutical-industry researchers in a broader number of aspects of the drug discovery and development continuum, including structural biology and translational research. This collaborative effort has played a role in molecularly targeted drugs such as cetuximab, trastuzumab, imatinib, and new promising drug candidates such as OSI-906. Cooperative efforts by industry and academia have also provided important insights to optimize the use of such agents in the clinic. This review aims to emphasize the need for academic/industrial collaborations for success and efficiency through the drug discovery and development continuum, and will highlight several examples of collaborations between academic and industrial scientists that facilitated the development of molecularly targeted antitumor agents into the clinic. ***Mt Sinai J Med 77:358–365, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: biotechnology, molecularly targeted drugs, structural biology, translational research, tumor biology.

Address Correspondence to:

Elizabeth Buck
OSI Pharmaceuticals
Farmingdale, NY
Email: ebuck@osip.com

Historically, discovery and development of cancer drugs has been the domain of large pharmaceutical companies. Academic institutions would become involved in the extensive clinical trials required to show proof of efficacy and clinical benefit of new cancer drugs. Table 1 shows the various steps and activities involved in drug discovery and development, along with which activities were mainly conducted by academia or the pharmaceutical industry. Over the past 2 decades, our knowledge of the molecular pathways that drive tumor-cell proliferation and survival has spurred the discovery and development of molecularly targeted therapeutics against a number of previously unassailable types of tumors. Thus emerged the promise of a new generation of molecularly targeted therapies that were more effective and less toxic than the traditional cytotoxic chemotherapeutics. Both biologic and small-molecule inhibitors directed toward either mutant forms of proteins expressed specifically by tumor cells or toward proteins to which tumor cells have become oncoaddicted

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20195

© 2010 Mount Sinai School of Medicine

Table 1. *Historical Roles of Academia and Industry in Drug Discovery and Development.*

	Target Identification	Target Validation	Lead Identification	Lead Optimization	Pre-IND Studies	Clinical Studies	Commercialize
Activities	Basic research, publications, databases	Basic research, publications, genetic models, cell models	Assay development, compound libraries, virtual libraries, screening, rational design	Medicinal chemistry, molecular modeling, refine potency, refine selectivity, refine pharmacology	Candidate development, toxicology, pharmacology, process chemistry, define efficacy, ID biomarkers	Phase I, Phase II, Phase III	Launch, marketing, sales
Academia	X	X				X	
Industry			X	X	X	X	X

Abbreviations: ID, identify; IND, Investigational New Drug Application.

have been developed and clinically approved for the treatment of several cancers. The desire to target new classes of proteins and improve the output of drug discovery has spurred advances in rational structure-based drug design and has fostered collaborations between industry academic scientists with complementary areas of expertise in a variety of biology and chemistry disciplines. With our greater understanding of the underlying molecular mechanisms of oncogenic transformation and tumorigenicity, we have gained a deeper appreciation for the inherent heterogeneity that exists within individual cancer types or histologies. Identifying molecular biomarkers that can be used to identify those patient subsets most likely to benefit from a specific drug has become a major goal in this new era of personalized or tailored therapy. In addition to tumor heterogeneity, we have also gained an understanding of the signaling plasticity that exists within tumor cells, allowing tumors to circumvent a targeted agent through either the acquisition of a drug-resistance mutation or the activation of an alternate signaling pathway to maintain tumor-cell growth and survival. This has emphasized the importance of translational research to enable personalized anticancer therapeutic strategies aimed at delivering the right drug(s) to the right patients at the right time and with the right dose. Many innovations for molecularly targeted anticancer therapeutics have been driven by collaborative efforts between pharmaceutical companies and academic researchers. This review will emphasize the continued need for collaborative industry-academia efforts, which can synergize in drug discovery and development. Four examples will be highlighted–cetuximab, trastuzumab, imatinib and dasatinib, and OSI-906–where efforts from both industry and academic scientists synergized to enable drug-discovery achievements and the delivery of anticancer therapeutics to the clinic. These examples provide a model for future collaborative efforts.

Many innovations for molecularly targeted anticancer therapeutics have been driven by collaborative efforts between pharmaceutical companies and academic researchers. This review will emphasize the continued need for collaborative industry-academia efforts, which can synergize in drug discovery and development.

DEVELOPMENT OF CETUXIMAB

Academic research over the past 4 decades has shown that inappropriate or overactive signaling by the epidermal growth factor receptor (EGFR) is an important contributor to specific cancers, especially non–small-cell lung cancer (NSCLC), squamous cell cancer of the head and neck (SCCHN), and colorectal cancer (CRC). Many tumor cell lines overexpress EGFR and/or overexpress the ligands for EGFR.[1] Nonmalignant cells ectopically overexpressing EGFR can be oncogenically transformed.[2,3] Neutralizing antibodies that bind to the extracellular domain of EGFR and block ligand binding are potent inhibitors of proliferation of tumor cell lines overexpressing EGFR.[3–5] Subsets of several cancer types, including NSCLC and SCCHN, overexpress EGFR.[1,3] To investigate the role of EGFR in human cancers and whether targeting EGFR offered a strategy for therapeutic intervention, John Mendelsohn and his colleagues in academia developed several monoclonal antibodies against the extracellular domain of EGFR.[5] Some of these EGFR-targeted monoclonal antibodies were shown to effectively inhibit the signaling activity of EGFR and, more important, resulted in significant inhibition of tumor growth and survival. Hybritech, Inc. licensed one of Mendelsohn's mouse monoclonal antibodies (mAb), mAb 225, from the University of California and embarked on scale-up production, formulation, and preclinical pharmacology.[4] A phase I clinical trial at Memorial Sloan-Kettering Cancer Center showed that mAb 225 was safe and well tolerated in humans, showed good pharmacokinetic properties, and selectively localized at tumor sites. All of the patients, however, developed human anti-mouse antibodies (HAMA). To overcome the HAMA reaction to mAb 225, the National Cancer Institute Biologics Decision Network Committee issued a contract for a research organization to develop a humanized human-mouse chimeric version of mAb 225, which became known as C225. Though Hybritech was involved in the initial commercialization effort of Mendelsohn's antibody, the rights to C225 were returned to the University of California following Hybritech's acquisition by Eli Lilly. In 1993, ImClone obtained a license from the University of California for the clinical development and commercialization of C225. ImClone scaled up manufacture of C225, which became known as cetuximab (Erbitux), and sponsored clinical trials with academic collaborators to evaluate the efficacy of cetuximab in different types of cancer patients. The results from a clinical trial on CRC patients with academic collaborators led to US Food and Drug Administration approval in 2004 for cetuximab in combination with irinotecan for the treatment of patients with EGFR-expressing, irinotecan-refractory CRC.[4]

The background science and rationale for EGFR as a target for cancer drug discovery and the key tool reagent, mAb 225, which was developed into a clinical therapeutic, were the result of the efforts by Mendelsohn and many other academic investigators. Manufacture and clinical development were mainly driven by industry efforts, initially by Hybritech and then by ImClone, through sponsorship of clinical trials with academic groups. Though cetuximab is an approved and marketed drug, not all treated patients benefit. This has led to efforts to identify biomarkers and accompanying diagnostic tests to identify the patients most likely to benefit from treatment with cetuximab. EGFR expression has been extensively studied, but has not been shown to be particularly useful as a predictive biomarker.[6] The *KRAS* mutation status appears to predict a poor response to cetuximab in CRC cancer patients, but may not be predictive in NSCLC patients.[7] This suggests that mechanistically, *KRAS* mutations alone may not preclude cetuximab from being effective, but that other tumor-type-specific factors are involved and need to be considered in developing future diagnostic biomarkers. These academic observations have major implications for the use and commercialization of cetuximab, therefore an area for mutual academic and industry study. Cetuximab is currently marketed by Bristol-Myers Squibb/ImClone and approved for certain types of CRC and SCCHN patients in specified single-agent and combination modalities.

DEVELOPMENT OF HERCEPTIN

The oncogenic properties of the *neu* gene in rats were first described by Robert Weinberg's laboratory at the Massachusetts Institute of Technology.[8] Several groups, including scientists at Genentech, identified the human homolog of *neu*, *HER2*.[9] Subsequent studies by other academic investigators showed that the protein tyrosine kinase activity of HER2 could oncogenically transform normal cells, either through activating mutations or overexpression of the wild-type protein.[1,9] Drs Slamon and McGuire in academia and Dr Ulrich in industry at Genentech noted the clinical relevance of HER2 in breast cancer[9] and embarked on an oncology collaboration targeting HER2. Scientists at Genentech developed mAbs against human the extracellular domain of HER2 and were able to show that some of the mAbs were potent growth inhibitors of HER2-expressing tumor cells.[10] The most promising mAb, muMAb4D5, was evaluated in a phase I study to assess safety, pharmacokinetics, and tumor localization. This mAb against HER2 was well tolerated and localized in tumors expressing HER2 protein, though patients did develop a HAMA response. MAb4D5 was humanized by Genentech[9] to minimize the emergence of a HAMA response, and was provided to academic laboratories for extensive preclinical characterization in tumor cell lines as a single agent and in combinations with chemotherapeutic agents. The results from the preclinical studies supported clinical evaluation of the humanized MAb4D5, now known as trastuzumab (Herceptin). Successful clinical studies involving Genentech and academic investigators resulted in the FDA approval of Herceptin for patients with certain types of breast cancer in 1998.[11]

Though the rationale for HER2/neu as a target for cancer drug discovery occurred through research conducted in academic laboratories, Genentech was the industry driver developing a mAb that served as an important tool providing proof-of-concept data as well as insights into how mAbs could be used in the clinic against HER2. Through academic collaborations, Genentech was able to generate preclinical supportive data as well as the clinical studies that resulted in Herceptin's FDA approval.[11] Though trastuzumab is effective in some HER2-overexpressing breast cancer patients, the biomarker determination of HER2 and the underlying mechanism of action is an active area of study for both academia and industry. This includes which methodology (eg, immunohistochemistry versus fluorescent in situ hybridization) to use in determining HER2 status to select patients for treatment with trastuzumab.[12] Trastuzumab is a chimeric humanized mouse antibody currently approved for patients with certain types of breast cancer in specified single-agent and combination modalities and is marketed by Genentech/Roche.

TARGETING BCR-ABL FOR THE TREATMENT OF CHRONIC MYELOGENOUS LEUKEMIA

The Philadelphia (Ph) chromosome is derived from a translocation involving regions of chromosomes 9 and 22, creating a fusion of the breakpoint cluster region (BCR) gene with the gene encoding the ABL, a SRC-related tyrosine kinase important in cellular proliferation and differentiation. The resulting BCR-ABL fusion protein is constitutively activated, leading to cellular transformation, and it is responsible for the development of most chronic myelogenous leukemias (CML).[13] Small molecules targeting BCR-ABL have been clinically approved for the treatment of Ph^+ CML, and the discovery and development of these agents was enabled in large part by work conducted in concert by academic institutions and pharmaceutical companies. High-throughput screening efforts at Novartis Pharmaceuticals led to the identification of a small-molecule ATP-competitive inhibitor of BCR-ABL, STI-571 (imatinib, Gleevec/Glivec), which was approved in 2001 for Ph^+ CML. Imatinib is the first example of an anticancer agent that specifically targets a mutant protein expressed by tumor cells and responsible for aberrant cell proliferation. Because it specifically targets tumor cells, imatinib has achieved complete remissions in CML without significant toxicities.[14] Although most CML patients initially respond to imatinib, acquired resistance to this agent occurs in select patients through long-term use. Studies conducted at academic institutions revealed the molecular mechanism for resistance, which ultimately enabled the discovery and development of next-generation therapeutics for imatinib-resistant CML. Dr Charles Sawyers at the University of California, Los Angeles, first identified the point mutation within the kinase domain of BCR-ABL that conferred resistance to imatinib therapy, and subsequently a collaboration between Dr Sawyers and Dr John Kuriyan at University of California Berkeley led to the identification of 15 distinct mutations that occurred in patients with acquired imatinib resistance.[15–17] Analysis of a crystal structure solved by Dr Kuriyan's group provided structural insights into the molecular mechanism by which these mutations conferred resistance. Although some of these mutations were involved in the direct binding of imatinib to the ATP binding site, many resided well outside of direct contact with imatinib. Dr Kuriyan's structure revealed that imatinib bound to BCR-ABL only when the activation loop was positioned in the closed conformation. Mutations in imatinib-resistant tumors blocked the ability of BCR-ABL to adopt the closed conformation, resulting in weak binding of imatinib to the protein. This understanding of the structural mechanism for imatinib resistance formed the basis for the next generation of BCR-ABL inhibitors. Since the imatinib-resistance mutations prevented BCR-ABL from adopting the closed inactive conformation, the investigators postulated that these mutant forms of BCR-ABL may bind better to ATP-competitive inhibitors without preference for the open or closed conformation of the enzyme. Dr Brian Druker's group at the Oregon Health & Science University Knight Cancer Institute found that the Src family kinase inhibitor PD180970, which binds preferentially to the active conformation of the kinase, was efficacious against BCR-ABL variants harboring clinically relevant imatinib-resistance mutations.[18] Although PD180970 lacked suitable drug-like properties, it provided important proof of concept for targeting the active conformation of BCR-ABL upon imatinib resistance. This understanding spawned a collaboration between investigators at Bristol-Myers Squibb Pharmaceuticals and Dr Sawyers leading to the identification and ultimate clinical approval of a compound, BMS-354825 (dasatinib), which exhibited efficacy in CML patients whose cancers had progressed on imatinib treatment.[19–23] Dasatinib inhibits ABL and other Src family kinases regardless of open or closed conformations. As predicted, investigators found that dasatinib inhibited both wild-type and mutant ABL kinase, and exhibited antitumor efficacy in vivo for CML tumors harboring imatinib-resistant BCR-ABL. A 2006 study initiated by Bristol-Myers Squibb demonstrated that dasatinib achieved hematological response in the majority of CML patients who had progressed on imatinib.[23]

The discovery and development of BCR-ABL kinase inhibitors for the treatment of CML has been enabled through an iterative cycle of studies conducted by scientists and clinicians in both academia and industry. This story illustrates how industry provided important know-how in high-throughput screening and medicinal chemistry to generate lead drug candidates, and work in academic laboratories with expertise in kinase structure and function provided important insights into structural mechanisms for resistance, leading to the development of next-generation inhibitors with improved resistance profiles.

> *In the discovery and development of breakpoint cluster region (BCR)-ABL kinase inhibitors, industry provided important know-how in high-throughput screening and medicinal chemistry to generate lead drug candidates, and academic laboratories with expertise in kinase structure and function provided important insights into structural mechanisms for resistance, leading to the development of next-generation inhibitors with improved resistance profiles.*

DISCOVERY OF OSI-906

The receptor for insulin-like growth factor (IGF-1R) is a major determinant of both embryonic development and postnatal growth. Embryos from mice null for IGF-1R are less than half the weight of their wild-type littermates.[24] IGF-1R is also a critical determinant of tumor-cell growth and survival. Work in Dr Renato Baserga's laboratory demonstrated that IGF-1R is required for transformation for a number of oncogenes including *RAS*, and IGF-1R promotes tumor growth in vivo.[25–27] The IGF-1R signaling pathway is also associated with acquired resistance toward a number of different antitumor agents including cytotoxic chemotherapeutics and radiation as well as other molecularly targeted agents. A growing body of data indicates that the insulin receptor (IR) also mediates tumor-cell growth and survival.[28–30] Given the important role of IGF-1R and IR in cancer, these receptors have been intensely targeted for the design of anticancer therapeutics, and both biologic and small-molecule tyrosine kinase inhibitors of IGF-1R and/or IR are currently in advanced clinical development. Humanized mAbs that specifically neutralize IGF-1R activity are being explored in the clinic; however, small molecules that dually inhibit IGF-1R and IR may exhibit greater antitumor activity for tumors where both receptors are coexpressed and activated. OSI-906 is a small-molecule dual IGF-1R/IR inhibitor currently in advanced clinical study.[31] OSI-906 exhibits potent inhibition of both IGF-1R and IR, but does not exhibit activity when evaluated against an extensive panel of other kinases.

The discovery of this class of inhibitors began at OSI with various hit-to-lead identification techniques. These efforts ultimately led to the identification of a lead series of adenosine triphosphate (ATP) competitive IGF-1R/IR inhibitors containing a unique imidazopyrazine template substituted with a benzyoxyphenyl moiety.[31–33] These compounds were selective for IGF-1R and IR, but were unoptimized in respect to target potency as well as for drug metabolism and pharmacokinetic (DMPK) properties. In order to facilitate structure-based design efforts to complement ongoing lead-optimization medicinal chemistry efforts, OSI initiated a collaboration with Dr Todd Miller (Stony Brook University) and Dr Steven Hubbard (New York University) to determine the cocrystal structure for a lead compound of the benzyloxyphenyl-derived imidazopyrazine series with the catalytic domain of IR. Dr Hubbard had previously determined the crystal structures for these receptors and is an expert in understanding protein tyrosine kinase function at the structural level.

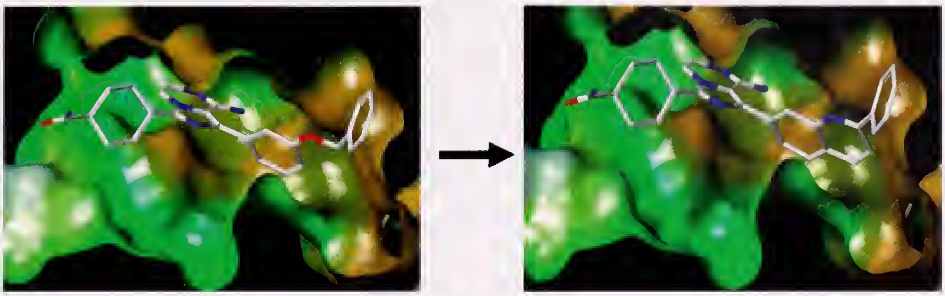


Fig 1. Co-crystal structure of lead OSI compounds bound to IR.

Dr Miller is an expert in the molecular mechanism for receptor tyrosine kinase signaling. Working together with OSI scientists, Dr Hubbard's and Dr Miller's groups determined the cocrystal structure of a lead OSI compound **1** and IR (Figure 1). Additionally, this cocrystal structure allowed for the construction of an IGF-1R homology model. Both cocrystal structure and homology modeling–based efforts, in conjunction with computer-assisted drug design (CADD) and medicinal chemistry efforts at OSI, led to the realization of key determinants of binding (Figure 2) for compounds within this series to both receptors that contributed to further maturation of the compound series. For example, critical hydrogen bonding interactions between compound and residues within the IR and IGF-1R catalytic domains were made evident, allowing insight into motifs that had to be preserved to maintain efficient binding. In addition to elucidating key binding determinants, 2 important structural observations were made: (1) the existence of an unoccupied hydrophobic pocket adjacent to the molecule, and (2) the coplanar nature and bioactive conformation of the benzyloxyphenyl pharmacophore. Based on these observations, a rational structure-based drug-design approach ensued at OSI, leading to the replacement of the benzyloxyphenyl moiety with 2-phenyl-quinolin-7-yl. The resulting molecule exhibited a 20-fold improvement in receptor binding as well as improved DMPK properties.

A subsequent IGF-1R crystal structure in collaboration with Drs. Miller and Hubbard was determined for an advanced quinolinyl lead molecule PQIP. This structure confirmed the predicted binding mode and validated the IGF-1R homology model. Further medicinal chemistry refinement of the series at OSI ultimately led to a best-in-series dual IGF-1R/IR inhibitor, OSI-906 (Figure 3). Paralleling in vitro autophosphorylation assays in the Miller lab allowed realization that compounds such as PQIP and OSI-906 bind to both the unphosphorylated "inactive" receptor as well as the tris-phosphorylated "activated" receptor. This binding characteristic is unique compared with other receptor tyrosine kinases that are currently clinically approved, including lapatinib (binding only to the inactive conformation) and erlotinib (binding only to the active conformation). This distinguishing binding characteristic is linked to the atypical binding mode of this class of compounds, which upon binding to the ATP pocket induce the

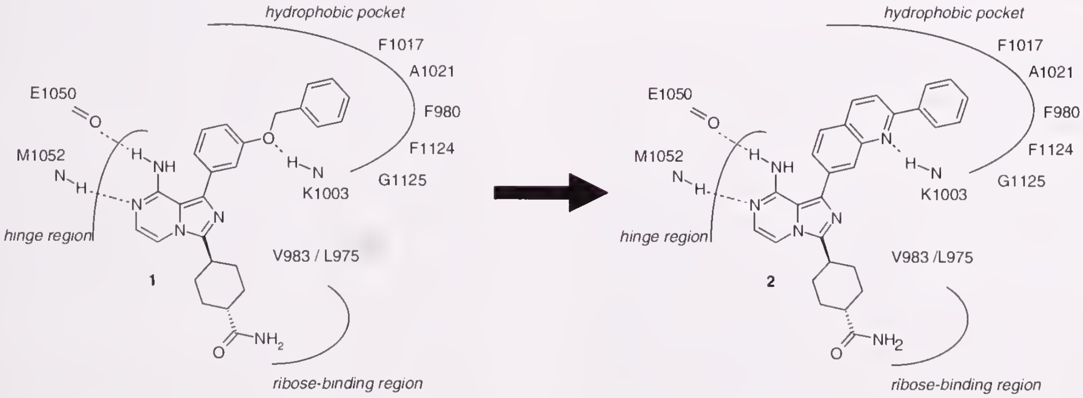


Fig 2. Structure of compound **1**, highlighting key contact regions with receptor that were realized following analysis of a co-crystal structure with receptor and computational modeling. Understanding these pharmacophores led to the rational structure based design of OSI compound **2**, exhibiting improved receptor binding properties.

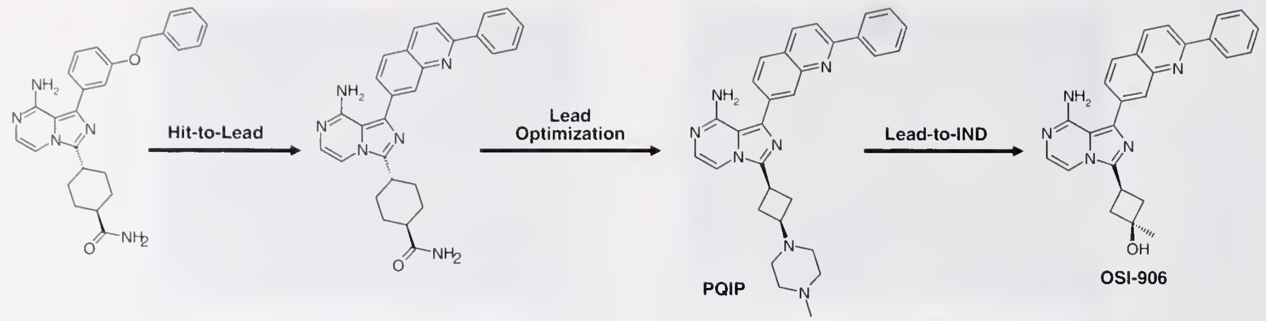


Fig 3. Schematic illustrating the progression of a lead IGF-1R inhibitor series at OSI, culminating in the IND compound OSI-906, which is currently in advanced clinical development.

expulsion of the activation loop from the active site and force the loop to adopt an extended conformation that is neither autoinhibitory nor active. This unique binding mode is proposed to explain the high degree of selectivity noted with this series of dual IGF-1R/IR inhibitors. The iterative collaboration between OSI Pharmaceuticals and researchers at 2 neighboring universities illustrates how industrial and academic resources can be leveraged in the discovery of antitumor therapeutics.

> *The iterative collaboration between OSI Pharmaceuticals and researchers at 2 neighboring universities illustrates how industrial and academic resources can be leveraged in the discovery of antitumor therapeutics.*

CONCLUSION

In the past, cancer target identification for therapeutic intervention and advanced clinical studies were the domain of academia. Drug discovery, development, and diagnostics were primarily the domain of industry. Recently, industry has taken a more active role in target identification and development, as exemplified by trastuzumab, and academia has demonstrated that it can be actively engaged in drug development, as exemplified by cetuximab. The discovery and development of OSI-906 and imatinib illustrates how industry and academia have worked together throughout many aspects and phases of drug discovery and development. Industry-academia collaborations continue to evolve in the new era of biomarker discovery and validation to support the development of truly personalized therapies. These examples serve as models that can be used to guide future drug discovery and development efforts, where collaborations between industry and academia accelerate both scientific innovation and the delivery of improved therapies to patients suffering from devastating diseases like cancer.

DISCLOSURES

Potential conflict of interest: EB, MM, and KI are employees of OSI Pharmaceuticals.

REFERENCES

1. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2001; 2: 127–137.
2. Arteaga CL. The epidermal growth factor receptor: from mutant oncogene in nonhuman cancers to therapeutic target in human neoplasia. *J Clin Oncol* 2001; 19(suppl 18): 32–40.
3. Mendelsohn J, Baselga J. The EGF receptor family as targets for cancer therapy. *Oncogene* 2000; 19: 6550–6565.
4. Mendelsohn J. Development of a targeted treatment for cancer: the example of C225 (Cetuximab). In: Kaufman HL, Wadler S, Antman K, eds. *Cancer Drug Discovery and Development: Molecular Targeting in Oncology*. Totowa, NJ: Humana Press; 2008; 559–571.
5. Sato JD, Kawamoto T, Le AD, et al. Biological effects in vitro of monoclonal antibodies to human epidermal growth factor receptors. *Mol Biol Med* 1983; 1: 511–529.
6. Chung KY, Shia J, Kemeny NE, et al. Cetuximab shows activity in colorectal cancer patients with tumors that do not express the epidermal growth factor receptor by immunohistochemistry. *J Clin Oncol* 2005; 23: 1803–1810.
7. Bardelli A, Siena S. Molecular mechanisms of resistance to cetuximab and panitumumab in colorectal cancer. *J Clin Oncol* 2010; 28: 1254–1261.
8. Schechter AL, Stern DF, Vaidyanathan L, et al. The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen. *Nature* 1984; 312: 513–516.

9. Carter P, Fendly BM, Lewis GD, Sliwkowski MX. Development of herceptin. *Breast Dis* 2000; 11: 103–111.
10. Fendly BM, Winget M, Hudziak RM, et al. Characterization of murine monoclonal antibodies reactive to either the human epidermal growth factor receptor or HER2/neu gene product. *Cancer Res* 1990; 50: 1550–1558.
11. Shepard HM, Jin P, Slamon DJ, et al. Herceptin. *Handb Exp Pharmacol* 2008; 183–219.
12. Sauter G, Lee J, Bartlett JM, et al. Guidelines for human epidermal growth factor receptor 2 testing: biologic and methodologic considerations. *J Clin Oncol* 2009; 27: 1323–1333.
13. Mauro MJ, Druker BJ. Chronic myelogenous leukemia. *Curr Opin Oncol* 2001; 13: 3–7.
14. Druker B. Signal transduction inhibition: results from phase I clinical trials in chronic myeloid leukemia. *Semin Hematol* 2001; 38(3 suppl 8): 9–14.
15. Gorre ME, Mohammed M, Ellwood K, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* 2001; 293(5531): 876–880.
16. Gorre ME, Sawyers CL. Molecular mechanisms of resistance to STI571 in chronic myeloid leukemia. *Curr Opin Hematol* 2002; 9: 303–307.
17. Shah NP, Nicoll JM, Nagar B, et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2002; 2: 117–125.
18. La Rosée P, Corbin AS, Stoffregen EP, et al. Activity of the Bcr-Abl kinase inhibitor PD180970 against clinically relevant Bcr-Abl isoforms that cause resistance to imatinib mesylate (Gleevec, STI571). *Cancer Res* 2002; 62: 7149–7153.
19. Burgess MR, Skaggs BJ, Shah NP, et al. Comparative analysis of two clinically active BCR-ABL kinase inhibitors reveals the role of conformation-specific binding in resistance. *Proc Natl Acad Sci U S A* 2005; 102: 3395–3400.
20. Druker BJ. Circumventing resistance to kinase-inhibitor therapy. *N Engl J Med* 2006; 354: 2594–2596.
21. Guilhot F, Apperley J, Kim DW, et al. Dasatinib induces significant hematologic and cytogenetic responses in patients with imatinib-resistant or -intolerant chronic myeloid leukemia in accelerated phase. *Blood* 2007; 109: 4143–4150.
22. Kujawski L, Talpaz M. Strategies for overcoming imatinib resistance in chronic myeloid leukemia. *Leuk Lymphoma* 2007; 48: 2310–2322.
23. Talpaz M, Shah NP, Kantarjian H, et al. Dasatinib in imatinib-resistant Philadelphia chromosome-positive leukemias. *N Engl J Med* 2006; 354: 2531–2541.
24. Louvi A, Accili D, Efstratiadis A. Growth-promoting interaction of IGF-II with the insulin receptor during mouse embryonic development. *Dev Biol* 1997; 189: 33–48.
25. Pollak M. Targeting insulin and insulin-like growth factor signalling in oncology. *Curr Opin Pharmacol* 2008; 8: 384–392.
26. Pollak M. Insulin and insulin-like growth factor signalling in neoplasia [published correction appears in *Nat Rev Cancer* 2009;9:224]. *Nat Rev Cancer* 2008; 8: 915–928.
27. Sell C, Rubini M, Rubin R, et al. Simian virus 40 large tumor antigen is unable to transform mouse embryonic fibroblasts lacking type 1 insulin-like growth factor receptor. *Proc Natl Acad Sci U S A* 1993; 90: 11217–11221.
28. Kalli KR, Falowo OI, Bale LK, et al. Functional insulin receptors on human epithelial ovarian carcinoma cells: implications for IGF-II mitogenic signaling. *Endocrinology* 2002; 143: 3259–3267.
29. Morrione A, Valentinis B, Xu SQ, et al. Insulin-like growth factor II stimulates cell proliferation through the insulin receptor. *Proc Natl Acad Sci U S A* 1997; 94: 3777–3782.
30. Vella V, Pandini G, Sciacca L, et al. A novel autocrine loop involving IGF-II and the insulin receptor isoform-A stimulates growth of thyroid cancer. *J Clin Endocrinol Metab* 2002; 87: 245–254.
31. Mulvihill MJ, Cooke A, Rosenfeld-Franklin M, et al. Discovery of OSI-906: a selective and orally efficacious dual inhibitor of the IGF-1 receptor and insulin receptor. *Future Med Chem* 2009; 1: 1153–1171.
32. Mulvihill MJ, Ji QS, Coate HR, et al. Novel 2-phenylquinolin-7-yl-derived imidazo[1,5-a]pyrazines as potent insulin-like growth factor-I receptor (IGF-IR) inhibitors. *Bioorg Med Chem* 2008; 16: 1359–1375.
33. Wu J, Li W, Craddock BP, et al. Small-molecule inhibition and activation-loop trans-phosphorylation of the IGF1 receptor. *EMBO J* 2008; 27: 1985–1994.

Drug-Based Therapies for Vascular Disease in Marfan Syndrome: From Mouse Models to Human Patients

Jason R. Cook, Harikiran Nistala, PhD,
and Francesco Ramirez, PhD

Mount Sinai School of Medicine, New York, NY

OUTLINE

Marfan Syndrome Overview
Fibrillin-1 Biology and Pathophysiology
Angiotensin Receptor Blocker Therapy
Future Research Directions and Therapeutic Modalities
Conclusion

ABSTRACT

Marfan syndrome is a congenital disorder of the connective tissue with a long history of clinical and basic science breakthroughs that have forged our understanding of vascular-disease pathogenesis. The biomedical importance of Marfan syndrome was recently underscored by the discovery that the underlying genetic lesion impairs both tissue integrity and transforming growth factor-β regulation of cell behavior. This discovery has led to the successful implementation of the first pharmacological intervention in a connective-tissue disorder otherwise incurable by either gene-based or stem cell–based therapeutic strategies. More generally, information gathered from the study of Marfan syndrome pathogenesis has the potential to improve the clinical management of common acquired disorders of connective-tissue degeneration. ***Mt Sinai J Med 77:366–373, 2010.*** © *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Francesco Ramirez
Department of Pharmacology
and Systems Therapeutics
Cardiovascular Institute Mount
Sinai School of Medicine
New York, NY
Email: francesco.ramirez@mssm.edu

Key Words: aortic aneurysm, connective tissue, fibrillin, losartan, Marfan syndrome, pharmacology, TGFβ, vascular therapy.

Marfan syndrome (MFS) is a common congenital disorder of the connective tissue with cardinal manifestations involving the skeletal, cardiovascular, and ocular systems.[1] The leading cause of morbidity and mortality in MFS is ascending aortic aneurysm that can precipitate acute dissection or rupture. For longer than a century, beginning in 1896 with Antoine Marfan's description of a young girl with unusual skeletal features, MFS has been the focus of intense clinical and basic science investigations that have led to a number of seminal discoveries and conceptual breakthroughs.[2] In 1956, Victor McKusick compiled the first comprehensive description of the MFS phenotype. McKusick further identified MFS as the prototype of a group of "heritable disorders of connective tissue" that he predicted to be accounted for by structural or metabolic dysfunctions of extracellular matrix (ECM) proteins.[3] Studies over the next 3 decades culminated in major clinical advances, such as the use of β-adrenergic blockade to reduce hemodynamic stress, the refinement of surgical interventions to replace the dilated aortic root, and the demonstration that mutations in the gene coding for the extracellular glycoprotein fibrillin-1 (*FBN1*) cause MFS.[4–7] Aside from validating McKusick's earlier prediction,

> *The leading cause of morbidity and mortality in Marfan syndrome (MFS) is ascending aortic aneurysm that can precipitate acute dissection or rupture.*

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20200

© 2010 Mount Sinai School of Medicine

identification of the MFS gene led to the development of molecular tools for prenatal diagnosis and of several mouse models of the disease. A decade-long characterization of these murine models of MFS established an unexpected correlation between mutations in fibrillin-1 and improper activation of transforming growth factor-β (TGFβ) signaling, which has in turn provided the evidence-based rationale for a more effective therapeutic strategy to blunt aortic aneurysm progression in MFS patients.[8] As such, the mouse findings have yielded important new insights into the clinical management of degenerative disorders of the connective tissue.

MARFAN SYNDROME OVERVIEW

MFS is inherited as an autosomal dominant trait with an estimated incidence of 1 per 5,000 individuals, of which ~25% are sporadic cases due to de novo *FBN1* mutations. The disease shows high penetrance and marked interfamilial and intrafamilial variability. Skeletal manifestations include disproportionate linear growth of digits, ribs, and limbs, craniofacial abnormalities, scoliosis, osteopenia, and joint hypermobility. Ocular disease in MFS includes dislocation of 1 or both lenses (ectopia lentis), strabismus, cataracts, and/or glaucoma.[9] Progressive enlargement of the ascending aorta and myxomatous valve changes are major cardiovascular manifestations that can cause chronic aortic regurgitation and aortic dissection. Additional abnormalities may involve the lungs, skin, integument, and muscle and adipose tissue.

MFS diagnosis is almost solely based on clinical findings and family history; molecular diagnosis is helpful in cases in which a familial *FBN1* mutation has been identified in affected family members. Guidelines established by the so-called Ghent nosology require a combination of manifestations in different organ systems that are defined as major and minor clinical criteria based on their frequency in the general population.[10] These diagnostic criteria help differentiate classic MFS from a number of closely related connective-tissue disorders that share some similarities *(FBN1* mutations) but differ in their repertoire of manifestations, natural history, and response to treatment. Differential diagnosis is particularly important for the timely management of cardiovascular complications in pediatric patients; however, extensive phenotypic variability and age-dependent onset of informative manifestations constitute significant challenges to MFS diagnosis in these individuals.

Differential diagnosis is particularly important for the timely management of cardiovascular complications in pediatric patients; however, extensive phenotypic variability and age-dependent onset of informative manifestations constitute significant challenges to MFS diagnosis in these individuals.

Vascular pathology in MFS is associated with changes in connective-tissue architecture manifested by reduced mechanical compliance, unbalanced ECM remodeling, and progressive aneurysm development that ultimately increase the risk of aortic-wall degradation.[11] These histopathological alterations, traditionally and improperly referred to as cystic medial necrosis, are also associated with the normal aging process, bicuspid aortic valve disease, and with increased hemodynamic stress in hypertension.[12,13] Until very recently, management of vascular disease in MFS has relied on routine echocardiograms to monitor aneurysm progression, administration of β-adrenergic blockers to reduce aortic-wall stress, and appropriately timed prophylactic surgery to prevent aortic complications.[14–16] Although these interventions have extended life expectancy from ~45 to ~72 years, alternative therapies are still needed to delay onset and progression of vascular disease, particularly in the most aggressive forms of MFS.[17,18] Additionally, strategies to manage associated comorbidities, such as musculoskeletal manifestations, are becoming increasingly important in the expanding population of aging MFS patients. However, the long-held belief that MFS is solely caused by loss of tissue integrity, together with the notion that fibrillin-1 assembly is largely confined to embryogenesis, represented a theoretically insurmountable impediment to the implementation of rational gene-based and/or cell-based therapies. This view has dramatically changed with recent findings that (1) fibrillin-1 may

Strategies to manage associated comorbidities, such as musculoskeletal manifestations, are becoming increasingly important in the expanding population of aging MFS patients.

be involved in elastic fiber maturation, homeostasis, and remodeling during postnatal life and, (2) the aortic architecture in fibrillin-1-mutant mice is normal at birth.[19] Specifically, murine models of MFS discussed in this review include *Fbn1*$^{mgR/mgR}$ mice, which underexpress fibrillin-1 and die of ruptured aortic aneurysm at ~2 months of age; *Fbn1*$^{C1039G/+}$ mice, which produce normal and structurally abnormal fibrillin-1 molecules and develop aortic aneurysm without dissection and rupture; and *Fbn1*$^{-/-}$ mice, which do not produce fibrillin-1 and die soon after birth due to the structural collapse of an improperly matured aortic wall.[20–22]

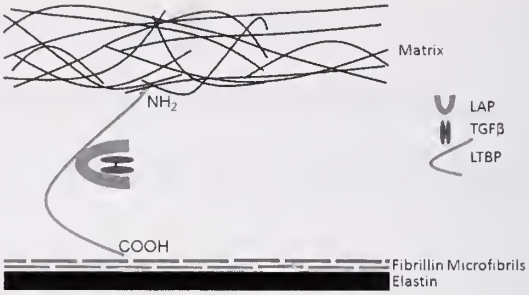


Fig 1. Fibrillin microfibrils modulate TGFβ incorporation into the extracellular matrix through binding to LTBP. **Abbreviations:** LAP, latency-associated peptide; LTBP, latent transforming growth factor β-binding protein; TGFβ, transforming growth factor β.

FIBRILLIN-1 BIOLOGY AND PATHOPHYSIOLOGY

FBN1 is a 350-kD cysteine-rich glycoprotein that polymerizes together with the structurally related fibrillin-2 into filamentous assemblies with an average diameter of 10 nm.[23] Fibrillin microfibrils impart key physical properties to virtually every organ system through the temporal and hierarchical assembly of tissue-specific macroaggregates through interactions with proteins such as elastin, proteoglycans, fibulins, and microfibril-associated glycoproteins.[23–28] In addition to its structural function, fibrillin microfibrils also control cell behavior through the sequestration of latent TGFβ complexes in the ECM via the dual interaction of latent transforming growth factor beta-binding protein (LTBP) with fibrillins and latent pro-TGFβ complexes (Figure 1).[27] TGFβ-1, -2, and -3 (hereafter collectively referred to as TGFβ) are multifunctional signaling molecules that orchestrate a large variety of cellular activities in a context-specific manner. TGFβ signals through receptor-induced stimulation of Smad-2 and Smad-3 (R-Smads) followed by the formation, nuclear translocation, and binding of activated R-Smad:Smad-4 complexes to specific DNA targets in association with transcriptional activators and/or repressors (Figure 2).[29,30] Inhibitory Smads, intracellular receptor trafficking, coreceptor activities, parallel activation of and cross-talk with mitogen-activated protein kinase (MAPK) signaling pathways, and interactions with ECM components constitute the multilayered regulatory system that modulates TGFβ bioavailability and controls the intensity and duration of TGFβ signaling.[31] Mouse models of MFS have strongly supported the notion that ECM sequestration imparts functional context to TGFβ activity by modulating the spatial and temporal release, as well as the concentration and presentation, of the ligands.[32]

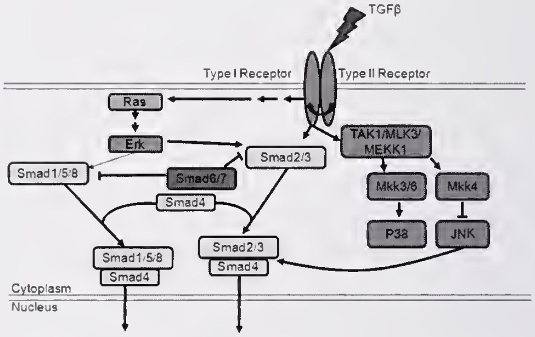


Fig 2. Simplified review of TGFβ signaling pathways. Molecules are color-coded according to their action: yellow, transcription factors; orange, kinases; red, inhibitory protein; gray, GTPase. **Abbreviations:** Erk, extracellular signal-regulated kinase; JNK, Jun NH$_2$–terminal kinase; MEKK1, MAPK/ERK-1 kinase; Mkk, MAP-kinase kinase; MLK3, mixed lineage kinase 3; TAK1, TGFβ-activated kinase; TGFβ, transforming growth factor β.

Early work has shown that ruptured aortic aneurysm in *Fbn1*$^{mgR/mgR}$ mice is preceded by a series of secondary cellular events that, in overlapping temporal succession, include medial calcification, excessive ECM accumulation and degradation, intimal hyperplasia, and adventitial inflammation.[22] Overall, these secondary events imply an unproductive tissue

> *Early work has shown that ruptured aortic aneurysm in fibrillin-1 (Fbn1)*$^{mgR/mgR}$ *mice is preceded by a series of secondary cellular events that, in overlapping temporal succession, include medial calcification, excessive extracellular matrix accumulation and degradation, intimal hyperplasia, and adventitial inflammation.*

remodeling/repair response that promotes intense elastolysis and that in turn exacerbates the collapse of a structurally compromised aortic wall. Subsequent studies implicated promiscuous TGFβ signaling as the molecular driver of improper tissue formation and remodeling in other *Fbn1* mutant mice.[8]

In addition to cardiovascular manifestations, MFS patients can present with pulmonary complications, including chronic obstructive lung disease and predisposition for pneumothorax.[33,34] Traditionally, this lung phenotype has been equated to destructive emphysema resulting from impaired tissue integrity.[1] Contrary to this prediction, however, widening of distal prealveolar saccules in newborn *Fbn1*-deficient mice occurs in the absence of inflammation or tissue destruction, suggesting a developmental defect as the underlying cause.[8] Upon closer inspection, the mouse phenotype was correlated with elevated TGFβ activity (as evidenced by greater nuclear accumulation of R-Smads) and increased apoptosis of lung cells, suggesting a developmental origin of the defect. Importantly, systemic administration of TGFβ-neutralizing antibodies during perinatal life rescued destructive emphysema in aged mutant mice, implying that early perturbations of tissue morphogenesis can predispose to late-onset degenerative processes. Similarly, increased cell proliferation and decreased apoptosis secondary to promiscuous TGFβ activity were associated with architectural alterations in the mitral valves of $Fbn1^{C1039G/+}$ mice.[35] Once again, systemic TGFβ antagonism during perinatal life corrected mitral valve abnormalities in $Fbn1^{C1039G/+}$ mice, in addition to improving aortic-wall architecture and enhancing regeneration of injured muscles.[36] Collectively, these findings established a causal correlation between impaired microfibril biogenesis and improper latent TGFβ activation. They also provided compelling evidence that pharmacological targeting of TGFβ signaling could be a more effective strategy than β-adrenergic blockade in preventing or restraining aortic aneurysm in MFS.

ANGIOTENSIN RECEPTOR BLOCKER THERAPY

The renin-angiotensin system (RAS) is essential for cardiovascular homeostasis. Angiotensin II (Ang II) binds two G-protein-coupled receptors (GPCR), angiotensin II type 1 receptor (AT_1) and angiotensin II type 2 receptor (AT_2), with opposing effects.[37] AT_1 activation by Ang II, autoantibodies, or mechanical stress can stimulate cell growth and migration, as well as inflammation and vasoconstriction.[38] Previous work has demonstrated highly localized expression of the AT_1 in the aortic arch to regions of nonlaminar flow.[39] Additional in vitro studies have shown that cardiomyocytes under mechanical stress exhibit increased AT_1 signaling in the absence of Ang II.[40] Activation of AT_1 signaling is the result of a mechanically induced conformational change of the receptor and can be inhibited by angiotensin receptor blockers (ARBs).[40] AT_2 activation can result in vasodilation, apoptosis, or inhibition of AT_1 signaling.[41] The differential outcomes of the Ang II receptors are probably best illustrated by the observation that AT_1 antagonism can prevent development of abdominal aortic aneurysm, whereas AT_2 antagonism accelerates it.[42] Experimentally induced renal and cardiac fibrosis in murine models has linked Ang II activation to TGFβ signaling through up-regulation of thrombospondin-1 (an activator of latent TGFβ), in addition to documenting the therapeutic efficacy of angiotensin-converting enzyme (ACE) inhibitors or ARBs in fibrotic disorders.[43–46]

Based on these correlative lines of evidence, the ARBs emerged as a promising therapeutic intervention for vascular therapy in MFS. The ARB losartan was therefore employed in a mouse preclinical trial to validate this hypothesis.[5] $Fbn1^{C1039G/+}$ mice were administered losartan, propanolol, or placebo beginning at embryonic day 14 or at 7 weeks of age for 10 or 6 months, respectively. Mice were sacrificed and aortic thickness and architecture were evaluated. Prenatal or postnatal administration of losartan, but not propanolol, prevented dilation of the aorta and destruction of aortic-wall architecture. Losartan therapy rescued aortic architecture to a greater extent than administration of TGFβ neutralizing antibody.[5] Similar findings were more recently obtained in a losartan preclinical trial that employed $Fbn1^{mgR/mgR}$ mice, thus reiterating drug efficacy in

> *In a mouse preclinical trial, losartan therapy rescued aortic architecture to a greater extent than administration of transforming growth factor β (TGFβ) neutralizing antibody. Then, a retrospective study of 18 pediatric patients with severe and rapidly progressive MFS who had been subject to ARB treatment for at least 1 year revealed a statistically appreciable decrease in the rate of aortic-root-diameter growth.*

curbing aneurysm progression in clinically distinct forms of MFS (our unpublished data). Importantly, the positive outcome of these mouse experiments provided proof-of-concept for the possible use of ARBs in treating MFS patients. A retrospective study of 18 pediatric patients with severe and rapidly progressive MFS who had been subject to ARB treatment for at least 1 year revealed a statistically appreciable decrease in the rate of aortic-root-diameter growth.[47] The positive results from both the mouse experiments and the retrospective study have resulted in a push toward better understanding the differences in the clinical efficacies of β-adrenergic, angiotensin-receptor blockade, and dual therapy. Currently there are multiple ongoing clinical randomized trials designed to assess the efficacy of the β-adrenergic blockers atenolol or nebivolol versus the ARB losartan in MFS individuals (clinical trials NCT00429364, NCT00593710, and NCT00683124).[48,49]

Despite encouraging evidence, there are practical and theoretical reasons to believe that some MFS patients might be unable to receive an ARB or respond fully to therapy. For example, ARBs are contraindicated in pregnancy, as inhibition of the fetal RAS can lead to fetopathy including renal failure, hypotension, oligohydramnios, and pulmonary hypoplasia.[50,51] Pregnant women represent a subset of the MFS patients that are at an increased risk for vascular complications secondary to increases in blood volume and heart rate associated with pregnancy.[52] MFS patients with aortic diameters >4 cm have a 10% risk of vascular complications during pregnancy.[53] Current therapies include β-blockers or elective surgical repair of the ascending aorta before conception to reduce the risk of complications.[52] In addition, long-term losartan treatment might exacerbate other morbid manifestations seen in MFS patients, including increased stiffness of the ascending aorta. Pulse pressure, defined as the difference between systolic and diastolic pressures, is influenced by arterial stiffness and has been shown to positively correlate with ascending aortic diameter.[54] Medical therapies that mitigate the arterial stiffness may further prevent vascular complications. Increased aortic stiffness has been observed in *Fbn1*$^{C1039G/+}$ mice in addition to decreased NO synthase activity, and treatment with losartan for 9 months was unable to mitigate the suppression of endothelial NO pathway.[55–57] Losartan, unlike other ARBs, is metabolized to an active metabolite EXP3174 through cytochrome P450 2C9 (CYP2C9) and has been previously demonstrated to have differential metabolism due to specific CYP2C9 polymorphisms.[58] It is theoretically possible that, under current dosing guidelines, MFS patients may appear refractory to losartan therapy due to polymorphisms in CYP2C9. Future studies should investigate whether orally active ARBs, such as telmisartan, are more effective in these populations.

> *Despite encouraging evidence, there are practical and theoretical reasons to believe that some MFS patients might be unable to receive an angiotensin receptor blocker (ARB) or respond fully to therapy.*

FUTURE RESEARCH DIRECTIONS AND THERAPEUTIC MODALITIES

The current model of MFS pathogenesis postulates that mutations affecting the structure or expression of fibrillin-1 preclude or decrease sequestration of latent TGFβ complexes in the ECM, thus rendering them more prone to or more accessible for activation. This view implies that physiological levels of latent TGFβ activators are sufficient to drive disease progression in MFS. An alternative and not mutually exclusive view is that a structurally abnormal ECM may also promote additional cellular responses that contribute to MFS pathogenesis by heightening latent TGFβ activation through the action of integrins, proteases, and/or other molecules.

In accordance with the above considerations, abnormally high p38 MAPK signaling has been reported to be an early determinant of greater R-Smad signaling in the aortas of *Fbn1*$^{-/-}$ mice, which mimic the neonatal lethal form of MFS (Figure 2).[19] This finding is consistent with the notion that TGFβ can also signal through Jun NH_2-terminal kinase (JNK), p38, and extracellular signal-regulated kinase (ERK1/2) MAPKs with different effects on cell performance.[59–63] Furthermore, MAPKs can be activated by a large number of environmental stimuli, including those mediated by AT_1, and MAPKs can regulate the expression of extracellular modulators of TGFβ signaling and influence R-Smad activity as well. Moreover, there is experimental evidence that Ang II can also induce the activity of ERK1/2 and JNK signaling.[64] It is therefore conceivable to argue that, depending on the nature of the *FBN1* mutation and on modifiers of stress responses and/or TGFβ activators, aortic aneurysm in MFS may progress by sequential activation of 1 or more MAPKs with discrete consequences for TGFβ signaling and cell performance. These complex molecular interactions may therefore account for the significant clinical variability of vascular disease in MFS.

Future strategies of MFS therapy should attempt to intercept and prevent disease pathogenesis at multiple levels. Aortic extracts from *Fbn1*$^{mgR/mgR}$ mice have been reported to induce macrophage chemotaxis, leading to a profound inflammatory response and activation of ECM-degrading and latent TGFβ-activating matrix metalloproteinases (MMPs).[65] Inability to sequester TGFβ through *FBN1* mutations in MFS coupled with increased MMP activation could exacerbate unopposed TGFβ signaling. Such a model of vascular pathogenesis would require the focused effort of losartan action on TGFβ signaling and other pharmacological means to mitigate the inflammatory response. Consistent with this proposed model, systemic MMP inhibition by doxycycline has been shown to improve aortic-wall architecture and delay ruptured aneurysm in *Fbn1*$^{C1039G/+}$ and *Fbn1*$^{mgR/mgR}$ mice, respectively.[66] Along the same lines, the preliminary finding that aortic architecture is improved to a greater extent using losartan as opposed to TGFβ-neutralizing antibody implicates the participation of factors other than improper TGFβ signaling in driving aneurysm progression in MFS. In this respect, it is worth noting that vascular manifestations similar to those of MFS are observed with mutations that affect smooth muscle–cell contraction, such as mutations in myosin heavy chain 11 (MYH11) and alpha-2 actin (ACTA2), or the plasma membrane receptor LRP1, an integrator of PDGF and TGFβ signals in the aortic wall.[67–69] Accordingly, it is possible that mutations in *FBN1* may negatively impact both ECM structural integrity and modulation of TGFβ bioavailability with multiple and complex consequences for aortic compliance and cell behavior. Ongoing investigations are therefore focusing on the mechanism of latent TGFβ activation and other potential contributors to aneurysm onset and progression with the intent of identifying new biological targets to improve vascular disease therapy in MFS.

> *Future strategies of MFS therapy should attempt to intercept and prevent disease pathogenesis at multiple levels.*

> *The preliminary finding that aortic architecture is improved to a greater extent using losartan as opposed to TGFβ-neutralizing antibody implicates the participation of factors other than improper TGFβ signaling in driving aneurysm progression in MFS.*

CONCLUSION

As in the past, MFS research has yielded new insights into the pathophysiology of human diseases. The new paradigm that ECM sequestration is a critical regulator of local TGFβ signaling has provided the conceptual framework of using traditional pharmacological means of therapy to counteract tissue degeneration in MFS and, conceivably, in age-related conditions affecting the integrity of connective tissues. Moreover, inappropriate TGFβ signaling in MFS has conceptualized the origin of clinical variability in this disease by identifying a number of candidate modifiers that are part of the regulatory network of TGFβ signaling. More generally, MFS represents a unique example of how characterization of monogenic disorders can inform and guide the clinical management of diseases in which gene-based or stem cell–based therapies are impracticable or impossible.

ACKNOWLEDGMENT

We thank Ms. Karen Johnson for help in preparing the manuscript.

Studies described in this review were supported by grants from the National Institutes of Health (AR-049 698) and from the National Marfan Foundation. Jason R. Cook is a trainee in the Integrated Pharmacological Sciences Training Program supported by grant T32GM062754 from the National Institute of General Medical Studies.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Judge DP, Dietz HC. Marfan's syndrome. *Lancet* 2005; 366: 1965–1976.
2. Marfan AB. Un cas de déformation congénitale des quatre membres plus pronocée aux extrémites charactérisée par l'allongement des os avec uncertain degréd' amincissement. *Bull Mem Soc Med Hop* 1896; 13: 220–226.

3. McKusick VA. Heritable disorders of connective tissue. III. The Marfan syndrome. *J Chronic Dis* 1955; 2: 609–644.
4. Dietz HC, Cutting GR, Pyeritz RE, et al. Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. *Nature* 1991; 352: 337–339.
5. Habashi JP, Judge DP, Holm TM, et al. Losartan, an AT1 antagonist, prevents aortic aneurysm in a mouse model of Marfan syndrome. *Science* 2006; 312: 117–121.
6. Gott VL, Greene PS, Alejo DE, et al. Replacement of the aortic root in patients with Marfan's syndrome. *N Engl J Med* 1999; 340: 1307–1313.
7. Salim MA, Alpert BS, Ward JC, Pyeritz RE. Effect of beta-adrenergic blockade on aortic root rate of dilation in the Marfan syndrome. *Am J Cardiol* 1994; 74: 629–633.
8. Neptune ER, Frischmeyer PA, Arking DE, et al. Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. *Nat Genet* 2003; 33: 407–411.
9. Maumenee IH. The eye in the Marfan syndrome. *Birth Defects Orig Artic Ser* 1982; 18: 515–524.
10. De Paepe A, Devereux RB, Dietz HC, et al. Revised diagnostic criteria for the Marfan syndrome. *Am J Med Genet* 1996; 62: 417–426.
11. Dean JC. Marfan syndrome: clinical diagnosis and management. *Eur J Hum Genet* 2007; 15: 724–733.
12. Carlson RG, Lillehei CW, Edwards JE. Cystic medial necrosis of the ascending aorta in relation to age and hypertension. *Am J Cardiol* 1970; 25: 411–415.
13. Cecconi M, Nistri S, Quarti A, et al. Aortic dilatation in patients with bicuspid aortic valve. *J Cardiovasc Med (Hagerstown)* 2006; 7: 11–20.
14. Pyeritz RE. Marfan syndrome: 30 years of research equals 30 years of additional life expectancy. *Heart* 2009; 95: 173–175.
15. Shores J, Berger KR, Murphy EA, Pyeritz RE. Progression of aortic dilatation and the benefit of long-term beta-adrenergic blockade in Marfan's syndrome. *N Engl J Med* 1994; 330: 1335–1341.
16. Milewicz DM, Dietz HC, Miller DC. Treatment of aortic disease in patients with Marfan syndrome. *Circulation* 2005; 111: e150–e157.
17. Silverman DI, Gray J, Roman MJ, et al. Family history of severe cardiovascular disease in Marfan syndrome is associated with increased aortic diameter and decreased survival. *J Am Coll Cardiol* 1995; 26: 1062–1067.
18. Murdoch JL, Walker BA, Halpern BL, et al. Life expectancy and causes of death in the Marfan syndrome. *N Engl J Med* 1972; 286: 804–808.
19. Pereira L, Andrikopoulos K, Tian J, et al. Targeting of the gene coding fibrillin-1 recapitulates the vascular phenotype of Marfan syndrome in the mouse. *Nat Genet* 1997; 17: 218–222.
20. Carta L, Pereira L, Arteaga-Solis E, et al. Fibrillins 1 and 2 perform partially overlapping functions during aortic development. *J Biol Chem* 2006; 281: 8016–8023.
21. Judge DP, Biery NJ, Keene DR, et al. Evidence for a critical contribution of haploinsufficiency in the complex pathogenesis of Marfan syndrome. *J Clin Invest* 2004; 114: 172–181.
22. Pereira L, Lee SY, Gayraud B, et al. Pathogenetic sequence for aneurysm revealed in mice underexpressing fibrillin-1. *Proc Natl Acad Sci U S A* 1999; 96: 3819–3823.
23. Sakai LY, Keene DR, Engvall E. Fibrillin, a new 350 kD glycoprotein, is a component of extracellular microfibrils. *J Cell Biol* 1986; 103: 2499–2509.
24. Ramirez F, Dietz HC. Fibrillin-rich microfibrils: structural determinants of morphogenetic and homeostatic events. *J Cell Physiol* 2007; 213: 326–330.
25. Kielty CM, Sherratt MJ, Marson A, Baldock C. Fibrillin microfibrils. *Adv Protein Chem* 2005; 70: 405–436.
26. Wagenseil JE, Mecham RP. New insights into elastic fiber assembly. *Birth Defects Res C Embryo Today* 2007; 81: 229–240.
27. Isogai Z, Ono RN, Ushiro S, et al. Latent transforming growth factor beta-binding protein 1 interacts with fibrillin and is a microfibril-associated protein. *J Biol Chem* 2003; 278: 2750–2757.
28. Ono RN, Sengle G, Charbonneau NL, et al. Latent transforming growth factor beta-binding proteins and fibulins compete for fibrillin-1 and exhibit exquisite specificities in binding sites. *J Biol Chem* 2009; 284: 16872–16881.
29. Massagué J, Gomis RR. The logic of TGFbeta signaling. *FEBS Lett* 2006; 580: 2811–2820.
30. Feng XH, Derynck R. Specificity and versatility in TGF-beta signaling through Smads. *Annu Rev Cell Dev Biol* 2005; 21: 659–693.
31. Moustakas A, Heldin CH. The regulation of TGFbeta signal transduction. *Development* 2009; 136: 3699–3714.
32. Ramirez F, Rifkin DB. Extracellular microfibrils: contextual platforms for TGFbeta and BMP signaling. *Curr Opin Cell Biol* 2009; 21: 616–622.
33. Robbesom AA, Koenders MM, Smits NC, et al. Aberrant fibrillin-1 expression in early emphysematous human lung: a proposed predisposition for emphysema. *Mod Pathol* 2008; 21: 297–307.
34. Bolande RP, Tucker AS. Pulmonary emphysema and other cardiorespiratory lesions as part of Marfan abiotrophy. *Pediatrics* 1964; 33: 356–366.
35. Ng CM, Cheng A, Myers LA, et al. TGF-beta-dependent pathogenesis of mitral valve prolapsed in a mouse model of Marfan syndrome. *J Clin Invest* 2004; 114: 1586–1592.
36. Cohn RD, van Erp C, Habashi JP, et al. Angiotensin II type 1 receptor blockade prevents TGF-beta-induced failure of muscle regeneration in multiple myopathic states [published correction appears in *Nat Med* 2007;13:511]. *Nat Med* 2007; 13: 204–210.
37. Gradman AH. Evolving understanding of the renin-angiotensin-aldosterone system: pathophysiology and targets for therapeutic intervention. *Am Heart J* 2009; 157(6 suppl): S1–S6.
38. Akazawa H, Yasuda N, Komuro I. Mechanisms and functions of agonist-independent activation in the angiotensin II type 1 receptor. *Mol Cell Endocrinol* 2009; 302: 140–147.
39. Ramkhelawon B, Vilar J, Rivas D, et al. Shear stress regulates angiotensin type 1 receptor expression in endothelial cells. *Circ Res* 2009; 105: 869–875.
40. Zou Y, Akazawa H, Qin Y, et al. Mechanical stress activates angiotensin II type 1 receptor without the involvement of angiotensin II. *Nat Cell Biol* 2004; 6: 499–506.
41. Berk BC. Angiotensin type 2 receptor (AT2R): a challenging twin. *Sci STKE* 2003; 181: PE16.
42. Daugherty A, Manning MW, Cassis LA. Antagonism of AT2 receptors augments angiotensin II-induced

abdominal aortic aneurysms and atherosclerosis. *Br J Pharmacol* 2001; 134: 865–870.

43. Lim DS, Lutucuta S, Bachireddy P, et al. Angiotensin II blockade reverses myocardial fibrosis in a transgenic mouse model of human hypertrophic cardiomyopathy. *Circulation* 2001; 103: 789–791.
44. Lavoie P, Robitaille G, Agharazii M, et al. Neutralization of transforming growth factor-beta attenuates hypertension and prevents renal injury in uremic rats. *J Hypertens* 2005; 23: 1895–1903.
45. Naito T, Masaki T, Nikolic-Paterson DJ, et al. Angiotensin II induces thrombospondin-1 production in human mesangial cells via p38 MAPK and JNK: a mechanism for activation of latent TGF-beta1. *Am J Physiol Renal Physiol* 2004; 286: F278–F287.
46. Seeland U, Schäffer A, Selejan S, et al. Effects of AT1- and beta-adrenergic receptor antagonists on TGF-beta1-induced fibrosis in transgenic mice. *Eur J Clin Invest* 2009; 39: 851–859.
47. Brooke BS, Habashi JP, Judge DP, et al. Angiotensin II blockade and aortic-root dilation in Marfan's syndrome. *N Engl J Med* 2008; 358: 2787–2795.
48. Lacro RV, Dietz HC, Wruck LM, et al. Rationale and design of a randomized clinical trial of beta-blocker therapy (atenolol) versus angiotensin II receptor blocker therapy (losartan) in individuals with Marfan syndrome. *Am Heart J* 2007; 154: 624–631.
49. Gambarin FI, Favalli V, Serio A, et al. Rationale and design of a trial evaluating the effects of losartan vs. nebivolol vs. the association of both on the progression of aortic root dilation in Marfan syndrome with FBN1 gene mutations. *J Cardiovasc Med (Hagerstown)* 2009; 10: 354–362.
50. Alwan S, Polifka JE, Friedman JM. Angiotensin II receptor antagonist treatment during pregnancy. *Birth Defects Res A Clin Mol Teratol* 2005; 73: 123–130.
51. Quan A. Fetopathy associated with exposure to angiotensin converting enzyme inhibitors and angiotensin receptor antagonists. *Early Hum Dev* 2006; 82: 23–28.
52. Goland S, Barakat M, Khatri N, Elkayam U. Pregnancy in Marfan Syndrome: maternal and fetal risk and recommendations for patient assessment and management. *Cardiol Rev* 2009; 17: 253–262.
53. Task Force on the Management of Cardiovascular Diseases During Pregnancy of the European Society of Cardiology. Expert consensus document on management of cardiovascular diseases during pregnancy. *Eur Heart J* 2003; 24: 761–781.
54. Jondeau G, Boutouyrie P, Lacolley P, et al. Central pulse pressure is a major determinant of ascending aorta dilation in Marfan syndrome. *Circulation* 1999; 99: 2677–2681.
55. Yang HH, Kim JM, Chum E, et al. Long-term effects of losartan on structure and function of the thoracic aorta in a mouse model of Marfan syndrome. *Br J Pharmacol* 2009; 158: 1503–1512.
56. Chung AW, Au Yeung K, Cortes SF, et al. Endothelial dysfunction and compromised eNOS/Akt signaling in the thoracic aorta during the progression of Marfan syndrome. *Br J Pharmacol* 2007; 150: 1075–1083.
57. Syyong HT, Chung AW, Yang HH, van Breemen C. Dysfunction of endothelial and smooth muscle cells in small arteries of a mouse model of Marfan syndrome. *Br J Pharmacol* 2009; 158: 1597–1608.
58. Yin T, Maekawa K, Kamide K, et al. Genetic variations of CYP2C9 in 724 Japanese individuals and their impact on the antihypertensive effects of losartan. *Hypertens Res* 2008; 31: 1549–1557.
59. Frey RS, Mulder KM. TGFbeta regulation of mitogen-activated protein kinases in human breast cancer cells. *Cancer Lett* 1997; 117: 41–50.
60. Atfi A, Djelloul S, Chastre E, et al. Evidence for a role of Rho-like GTPases and stress-activated protein kinase/c-Jun N-terminal kinase (SAPK/JNK) in transforming growth factor beta-mediated signaling. *J Biol Chem* 1997; 272: 1429–1432.
61. Mulder KM. Role of Ras and Mapks in TGFbeta signaling. *Cytokine Growth Factor Rev* 2000; 11: 23–35.
62. Hanafusa H, Ninomiya-Tsuji J, Masuyama N, et al. Involvement of the p38 mitogen-activated protein kinase pathway in transforming growth factor-beta-induced gene expression. *J Biol Chem* 1999; 274: 27161–27167.
63. Frey RS, Mulder KM. Involvement of extracellular signal-regulated kinase 2 and stress-activated protein kinase/Jun N-terminal kinase activation by transforming growth factor beta in the negative growth control of breast cancer cells. *Cancer Res* 1997; 57: 628–633.
64. Li X, Lee JW, Graves LM, Earp HS. Angiotensin II stimulates ERK via two pathways in epithelial cells: protein kinase C suppresses a G-protein coupled receptor-EGF receptor transactivation pathway. *EMBO J* 1998; 17: 2574–2583.
65. Guo G, Booms P, Halushka M, et al. Induction of macrophage chemotaxis by aortic extracts of the mgR Marfan mouse model and a GxxPG-containing fibrillin-1 fragment. *Circulation* 2006; 114: 1855–1862.
66. Xiong W, Knispel RA, Dietz HC, et al. Doxycycline delays aneurysm rupture in a mouse model of Marfan syndrome. *J Vasc Surg* 2008; 47: 166–172.
67. Guo DC, Pannu H, Tran-Fadulu V, et al. Mutations in smooth muscle alpha-actin (ACTA2) lead to thoracic aortic aneurysms and dissections [published correction appears in *Nat Genet* 2008;40: 255]. *Nat Genet* 2007; 39: 1488–1493.
68. Boucher P, Li WP, Matz RL, et al. LRP1 Functions as an atheroprotective integrator of TGFbeta and PDGF signals in the vascular wall: implications for Marfan Syndrome. *PLoS One* 2007; 2: e448.
69. El-Hamamsy I, Yacoub MH. Cellular and molecular mechanisms of thoracic aortic aneurysms. *Nat Rev Cardiol* 2009; 6: 771–786.

Role of Antibodies in Developing Drugs That Target G-Protein-Coupled Receptor Dimers

Chris Hipser, Ittai Bushlin, Achla Gupta, PhD, Ivone Gomes, PhD,
and Lakshmi A. Devi, PhD

Department of Pharmacology and Systems Therapeutics, Mount Sinai School of Medicine, New York, NY

OUTLINE

G-Protein-Coupled Receptor Dimerization and Allosterism
Antibodies as G-Protein-Coupled Receptor Drug Targets
Conclusion
Acknowledgments

ABSTRACT

G-protein-coupled receptors are important molecular targets in drug discovery. These receptors play a pivotal role in physiological signaling pathways and are targeted by nearly 50% of currently available drugs. Mounting evidence suggests that G-protein-coupled receptors form dimers, and various studies have shown that dimerization is necessary for receptor maturation, signaling, and trafficking. However, the physiological implications of dimerization in vivo have not been well explored because detection of GPCR dimers in endogenous systems has been a challenging task. One exciting new approach to this challenge is the generation of antibodies against specific G-protein-coupled receptor dimers. Such antibodies could be used as tools for characterization of heteromer-specific function; as reagents for their purification, tissue localization, and regulation in vivo; and as probes for mapping their functional domains. In addition, such antibodies could serve as alternative ligands for G-protein-coupled receptor heteromers. Thus, heteromer-specific antibodies represent novel tools for the exploration and manipulation of G-protein-coupled receptor-dimer pharmacology. ***Mt Sinai J Med 77:374–380, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: antibodies, dimers, GPCRs, G-protein-coupled receptors, heteromers.

Address Correspondence to:

Lakshmi A. Devi
Mount Sinai School of Medicine
New York, NY
Email: lakshmi.devi@mssm.edu

G-protein-coupled receptors (GPCRs) are a family of integral membrane proteins located on the cell surface whose activation induces second-messenger amplification. These receptors are drug targets in a wide range of major diseases such as asthma, cancer, diabetes, hypertension, obesity, chronic pain, and other neurological disorders.

> *Located on the cell surface, G-protein-coupled receptors (GPCRs) are a family of integral membrane proteins that are drug targets in a wide range of major diseases such as asthma, cancer, diabetes, hypertension, obesity, chronic pain, and other neurological disorders.*

For a number of years it was thought that individual GPCRs functioned as a monomeric unit. However, in recent years several studies showed that GPCRs can form dimers, such as homodimers and heteromers, as well as more complex oligomers.[1–3] There is also growing evidence suggesting that dimerization substantially modulates receptor function.[4] Studies have shown that receptor heteromers have important roles in receptor maturation, signaling, and trafficking.[5–8] Most of these studies were carried out in heterologous cell systems overexpressing different epitope-tagged GPCRs using a variety of assays

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20199

© 2010 Mount Sinai School of Medicine

including ligand binding, coimmunoprecipitation,[9,10] bioluminescence resonance energy transfer,[11–13] or fluorescence resonance energy transfer.[14–16] However, detection and characterization of specific GPCR dimers in an endogenous system has been challenging due to the lack of selective tools. We propose the generation of GPCR dimer–specific antibodies that could be used as tools for the characterization of heteromer pharmacology and signaling, and as reagents for the study of their in vivo localization and regulation. In addition, such antibodies could be developed as alternative highly selective ligands targeting GPCR heteromers (Figure 1); the functional and physiological diversity of GPCR dimers provides the necessary targets for enabling novel drug discovery.

In this review, we focus our attention on the potential implications of antibodies in the development of selective reagents targeting GPCR dimers.

It is now well established that GPCRs exist and function as dimers/oligomers. Interactions between identical protomers (a single GPCR) are referred to as homomers, and interactions between nonidentical protomers are referred to as heteromers. Additionally, there is growing evidence that heteromerization can generate receptors with novel characteristics, leading to altered pharmacological properties.

G-PROTEIN-COUPLED RECEPTOR DIMERIZATION AND ALLOSTERISM

It is now well established that GPCRs exist and function as dimers/oligomers. Interactions between identical protomers (a single GPCR) are referred to as homomers, and interactions between nonidentical protomers are referred to as heteromers. Additionally, there is growing evidence that heteromerization can generate receptors with novel characteristics, leading to altered pharmacological properties.[1,17,18] The most convincing evidence for receptor heteromerization has come from studies with metabotropic γ-aminobutyric acid$_B$ (GABA_B) receptors in heterologous systems where the functional receptor requires the coexpression of both the GABA_BR1 and GABA_BR2 subunits.[19] In the absence of GABA_BR2, the GABA_BR1 subunit is retained in the endoplasmic reticulum, whereas in the absence of GABA_BR1, the GABA_BR2 subunit is expressed at the cell surface but is nonfunctional; only the coexpression of the 2 subunits leads to plasma membrane expression of a functional receptor. Heteromerization leads to the masking of an endoplasmic reticulum retention sequence in GABA_{B1} by GABA_{B2}, allowing the heterodimers to be efficiently expressed at the cell surface.[20,21] This kind of association is found to occur very shortly after biosynthesis, meaning that dimerization could be key to receptor maturation in the case of this and other GPCRs.[7,21,22]

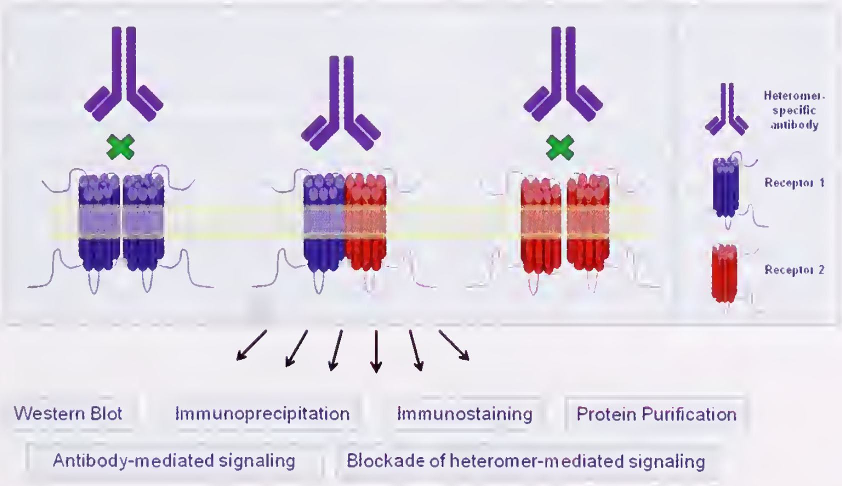

Heteromer-specific antibody
Receptor 1
Receptor 2
Western Blot
Immunoprecipitation
Immunostaining
Protein Purification
Antibody-mediated signaling
Blockade of heteromer-mediated signaling

Fig 1. A monoclonal antibody can be generated that specifically recognizes receptor heteromers but not homomers. It could be used as a tool to detect receptor heteromers in vivo and to characterize heteromer-specific signaling.

The concept that GPCR heteromerization could have a role in pharmacological diversity was first observed in studies on the δ-opioid receptors and κ-opioid receptors.[17] These studies show that the δ-κ complex exhibits ligand-binding, functional, and trafficking properties that are distinct from those of each individual receptor.[17] Studies also show that μ and δ receptors, once thought to only function individually, can dimerize and form a complex exhibiting distinct functional properties[23] that have profound effects on morphine-mediated analgesia. It is probable that additional pairings involving the μ-opioid receptor and other GPCRs could play crucial roles in regulating analgesia and addiction.[24] In another compelling case, the mammalian ability to perceive sweet tastes depends on the heteromerization of T1R2 and T1R3 taste receptors, whereas the umami taste relies on a T1R1/T1R3 dimer, indicating that heterodimerization between different subtypes of the taste receptor leads to modulation of ligand-binding properties resulting in the recognition of very distinctive tastes.[25] Studies also show that receptors thought to be normally involved in completely different physiological systems may form heterodimers,[26] thus allowing for a new level of specificity for drug targeting.

In some cases, GPCR dimerization/heterodimerization is associated with changes in ligand-binding affinity for the complex compared with the affinity for the homomer, implying heteromer-mediated modulation of the ligand-binding pocket. Several studies have shown positive or negative cooperativity in ligand binding following receptor coexpression likely resulting from receptor heterodimerization. Examples include GPCR dimer pairs such as muscarinic m2/m3,[27] somatostatin SSTR5/dopamine D2,[15] adenosine A2A/dopamine D1,[28] and the opioid δ/μ[23,29] receptors. These results suggest that heteromerization between pharmacologically different receptors provides a level of diversity that could generate new opportunities for the development of more selective compounds that would target specific heterodimers without affecting the activity of individual protomers.

In an attempt to develop drugs that target GPCR heterodimers, bivalent ligands have been developed.[30] Bivalent ligands are composed of 2 functional pharmacophores linked by a spacer that bind to the orthosteric sites of the protomers within a heteromer. In principle, bivalent ligands would preferentially bind to the heteromer; the latter is likely to be selectively expressed only in tissues coexpressing both of the interacting receptors, and this would lead to improved drug targeting and tissue selectivity. Some bivalent ligands were found to be effective in increasing the specificity of drug effects–for example, increased analgesia with a reduction in side effects such as development of tolerance–suggesting that bivalent ligands could be developed as potential novel analgesics.[31,32] A set of such ligands targeted both μ and δ receptors and consisted of a μ-agonist pharmacophore (oxymorphone) and a δ-antagonist pharmacophore (naltrindole; NTI) separated by spacers of variable lengths.[33] The antinociceptive activity of these bivalent ligands was reported to be better than that achieved by coadministration of the individual pharmacophores and could be titrated by manipulating the distance between the pharmacophores.[33] Another group of bivalent ligands targeting δ-κ heteromers was recently synthesized consisting of δ and κ antagonists (NTI and 5′-GNTI, respectively) linked by variable-length spacers.[34] One of these compounds (KDN-21) exhibited a much higher affinity and selectivity for δ-κ heterodimers relative to δ or κ receptors alone.[35] KDN-21 was found to have antinociceptive activity when administered directly into the spinal cord but not into the brain, supporting the notion that the relative levels of a heterodimer vary in different regions and that selectively targeting a heterodimer could increase the specificity of a drug with probable reduction of associated side effects. Taken together, these studies show that bivalent ligands to opioid receptors could be developed as novel analgesics with greater tissue selectivity and lesser degree of side effects normally associated with chronic opiate administration.

The possibility of allosterically modulating the function of individual protomers within a GPCR dimer makes the heteromer a new type of drug target. Classically, the majority of drugs target the receptor's orthosteric site, which is often the binding site for the endogenous agonist. Recent studies have begun to explore the possibility of targeting allosteric sites and have identified synthetic allosteric modulators (activator or inhibitors of GPCRs).[36,37] These alternate sites can bind small-molecule ligands, leading to the modulation of agonist binding to the orthosteric site. In the case of a GPCR dimer, a ligand bound to the orthosteric site of one protomer could function as an allosteric modulator of the other protomer within the pair, which could potentially alter the conformation and functional outcome of either protomer, receptor heteromer, and/or the whole complex. The design of such drugs would help us understand the unique pharmacology and signaling mechanisms activated by GPCR dimers given the mounting evidence suggesting that every GPCR dimer displays unique functional properties with different pharmacological characteristics.[15,27–29] It is likely that heteromerization produces a unique binding pocket that binds ligands with distinct affinity and displays unique ligand structure-activity relationships compared with receptor homomers. Thus it would be advantageous to begin to find smaller, drug-like synthetic compounds that would bind to allosteric sites in either protomer and alter the conformation of a dimer's orthosteric binding pocket. These smaller, single-target drugs might also be made to target specific sites in the novel binding site created in the pocket between the 2 interacting receptor proteins. In this context, dimer-selective antibodies would be useful agents for the exploration of dimerization within in vitro as well as in vivo systems.

One of the reasons why GPCR dimers are exciting drug targets is that a change in their expression levels could contribute to the development of disease symptoms in specific types of tissues.

One of the reasons why GPCR dimers are exciting drug targets is that a change in their expression levels could contribute to the development of disease symptoms in specific types of tissues. For example, it has been reported that preeclamptic hypertensive women exhibit a significant increase in heteromers between the AT_1 angiotensin receptor (which responds to the vasopressor angiotensin II) and the B_2 bradykinin receptor (which responds to the vasodepressor bradykinin). AT1 and B_2 bradykinin receptors normally regulate blood pressure by controlling vasoconstriction and smooth-muscle relaxation, respectively. During preeclampsia, the dimeric AT1-B_2 receptor complex becomes hypersensitive to angiotensin II and exhibits reduced affinity for the B_2 receptor agonist, bradykinin.[38] This example of in vivo alteration of GPCR pharmacology by dimerization indicates that dimers could be useful for controlling blood pressure. Disease-specific GPCR dimers are also thought to play roles in the regulation of cardiac muscle-cell function, asthma, schizophrenia, drug-related analgesia and tolerance, and other pathologies.[23,39–42] However, the lack of suitable tools to study the distribution and regulation of heteromers in vivo has made such studies difficult. Recent advances in antibody technology have begun to address some of these difficult questions; these are described below.

The lack of suitable tools to study the distribution and regulation of heteromers in vivo has made the study of disease-specific GPCR dimers difficult. Recent advances in antibody technology have begun to address some of these difficult questions.

ANTIBODIES AS G-PROTEIN-COUPLED RECEPTOR DRUG TARGETS

Antibodies have been used as tools for receptor characterization, purification, and localization and as probes for mapping their functional domains. Antibodies are now becoming integral tools in drug research and are even being developed as drugs. Their unique design makes them especially suited for attaining a high level of specificity for a large variety of organic, pharmacologically significant molecules and epitopes on larger molecules. In particular, monoclonal antibodies, which by definition recognize a single epitope, are quite useful in these applications, unlike polyclonal antibodies, which target multiple epitopes.

Antibodies are excellent diagnostic screening tools, as they can detect the domains involved in activity-mediated conformational changes of signaling proteins, including receptors.

Antibodies are excellent diagnostic screening tools, as they can detect the domains involved in activity-mediated conformational changes of signaling proteins, including receptors.[43–51] For example, a monoclonal antibody to the N-terminal region of rhodopsin exhibited a higher degree of recognition for activated receptors than for inactivated receptors even after detergent treatment, suggesting that photoactivation of rhodopsin induces a conformational change at the N-terminus that exposes an epitope that is recognized by the monoclonal antibody.[52] Recently, Gupta et al[53,54] showed that antibodies targeting the N-termini of family A GPCR homomers can discriminate between activation states of the receptors. Because the extent to which a certain antibody binds to a receptor can depend on whether the latter is activated by a functional ligand, an assay using antibodies could be used to effectively screen for novel GPCR ligands. This method recently led to the identification of hemopressin as an interesting new peptide antagonist of the CB1 cannabinoid receptor.[55] In addition, it will be important to develop heteromer-specific antibodies, given the critical role of receptor heteromers in certain cell-signaling and disease processes.

Several studies have shown that antibodies directed against GPCRs can act as allosteric receptor agonists or antagonists.[56–59]

> *Several studies have shown that antibodies directed against GPCRs can act as allosteric receptor agonists or antagonists.*

In the case of the δ opioid receptor, an antibody directed to the N-terminal region behaves like a classic agonist and activates the receptor.[58] For the β adrenergic receptor, a monoclonal antibody to the second extracellular loop had agonist-like activity, whereas a monovalent (Fab) fragment acted as an antagonist.[59] Antibodies are usually very effective when they target the extracellular N-termini and C-termini or the third intracellular loop of GPCRs.[53,60] Because these regions are highly diverse, the problem of antibodies binding to similarly structured molecules, which plagues synthetic drugs, is practically eliminated.

Antibodies can also be useful in examining the conformational changes of a particular protomer that arise upon binding of a ligand to the partner protomer.[57]

> *Antibodies can also be useful in examining the conformational changes of a particular protomer that arise upon binding of a ligand to the partner protomer.*

For example, an antibody binding to an allosteric site on one protomer could differentially detect whether 1 or both protomers are bound by selective orthosteric/allosteric ligands.[57] This could be used as a screening strategy for the detection of novel allosteric ligands targeting heteromers. Allosteric drugs could be useful because they leave the main active site free, only modulating its behavior. The side effects associated with modern orthosteric drugs are often due to the replacement of the receptor's endogenous ligand, altering homeostasis far more than is necessary. Thus, using antibodies as allosteric drugs could circumvent this problem.

Antibodies, including humanized antibodies, are currently being used to detect receptors, their activity states, and their location and density in the treatment of immune-system diseases and cancers.[61]

> *Antibodies, including humanized antibodies, are currently being used to detect receptors, their activity states, and their location and density in the treatment of immune-system diseases and cancers.*

As already mentioned, antibodies can act as functional modulators, direct agonists, inverse agonists, and antagonists. In addition, antibodies can be designed to carry marker molecules that bind to a selective secondary drug, or they may simply be designed to carry a toxin to, for example, kill target cancer cells.[62] This way, only cells expressing certain receptors or specific receptor dimers would be targeted and altered or destroyed. Currently, 22 monoclonal-antibody drugs have been approved by the US Food and Drug Administration for therapeutic use, and hundreds more are undergoing clinical trials. Other than vaccines, antibodies are the fastest-growing drug category and have the potential to bring under control a vast array of disorders.[62]

CONCLUSION

Mounting evidence suggests that GPCR dimerization is an important physiological method of restoring equilibrium. It is easy to see that dimer regulation is uniquely useful in the body and that understanding dimer pharmacology could lead to the development of far more selective drugs and to more effective treatments for diseases. Thus, antibodies against G-protein-coupled receptor dimers could be useful tools in detection of receptor dimers, in assessment of dimer pharmacology, and in screening during drug development. In the near future, antibodies could be generated and optimized to serve as drugs by directly targeting tissues that exhibit altered levels of GPCR dimers and selectively block or activate heteromer function.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants DA08862, DA19521, and 1P50GM071558-01A27398 (SBCNY) to Lakshmi A. Devi. Ittai Bushlin is a trainee in the Integrated Pharmacological Sciences Training Program supported by grant T32GM062754 from the National Institute of General Medical Sciences.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Bouvier M. Oligomerization of G-protein-coupled transmitter receptors. *Nat Rev Neurosci* 2001; 2: 274–286.
2. Angers S, Salahpour A, Bouvier M. Dimerization: an emerging concept for G protein-coupled receptor ontogeny and function. *Annu Rev Pharmacol Toxicol* 2002; 42: 409–435.
3. George SR, O'Dowd BF, Lee SP. G-protein-coupled receptor oligomerization and its potential for drug discovery. *Nat Rev Drug Discov* 2002; 1: 808–820.
4. Milligan G. G protein-coupled receptor dimerization: function and ligand pharmacology. *Mol Pharmacol* 2004; 66: 1–7.
5. Devi LA. G-protein-coupled receptor dimers in the lime light. *Trends Pharmacol Sci* 2000; 21: 324–326.
6. Jordan BA, Trapaidze N, Gomes I, et al. Dimerization of opioid receptors with beta 2-adrenergic receptors: a role in trafficking and mitogen activated protein kinase activation. *Proc Natl Acad Sci U S A* 2001; 98: 343–348.
7. Terrillon S, Bouvier M. Roles of G-protein-coupled receptor dimerization. *EMBO Rep* 2004; 5: 30–34.
8. Bulenger S, Marullo S, Bouvier M. Emerging role of homo- and heterodimerization in G-protein-coupled receptor biosynthesis and maturation. *Trends Pharmacol Sci* 2005; 26: 131–137.
9. Hebert TE, Moffett S, Morello JP, et al. A peptide derived from a beta2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation. *J Biol Chem* 1996; 271: 16384–16392.
10. Cvejic S, Devi LA. Dimerization of the delta opioid receptor: implication for a role in receptor internalization. *J Biol Chem* 1997; 272: 26959–26964.
11. Angers S, Salahpour A, Joly E, et al. Detection of beta 2-adrenergic receptor dimerization in living cells using bioluminescence resonance energy transfer (BRET). *Proc Natl Acad Sci U S A* 2000; 28: 97 3684–3689.
12. McVey M, Ramsay D, Kellett E, et al. Monitoring receptor oligomerization using time-resolved fluorescence resonance energy transfer and bioluminescence resonance energy transfer. The human delta-opioid receptor displays constitutive oligomerization at the cell surface, which is not regulated by receptor occupancy. *J Biol Chem* 2001; 276: 14092–14099.
13. Ayoub MA, Couturier C, Lucas-Meunier E, et al. Monitoring of ligand-independent dimerization and ligand-induced conformational changes of melatonin receptors in living cells by bioluminescence resonance energy transfer. *J Biol Chem* 2002; 277: 21522–21528.
14. Overton MC, Blumer KJ. G-protein-coupled receptors function as oligomers in vivo. *Curr Biol* 2000; 10: 341–344.
15. Rocheville M, Lange DC, Kumar U, et al. Receptors for dopamine and somatostatin: formation of hetero-oligomers with enhanced functional activity. *Science* 2000; 288: 154–157.
16. Dinger MC, Bader JE, Kobor AD, et al. Homodimerization of neuropeptide Y receptors investigated by fluorescence resonance energy transfer in living cells. *J Biol Chem* 2003; 278: 10562–10571.
17. Jordan BA, Devi LA. G-protein-coupled receptor heterodimerization modulates receptor function. *Nature* 1999; 399: 697–700.
18. Devi LA. Heterodimerization of G-protein-coupled receptors: pharmacology, signaling and trafficking. *Trends Pharmacol Sci* 2001; 22: 532–537.
19. Jones KA, Borowski B, Tamm JA, et al. GABA$_B$ receptor function as a heterotrimeric assembly of the subunits GABA$_B$R1 and GABA$_B$R2. *Nature* 1998; 396: 674–679.
20. Pin JP, Kniazeff J, Liu J, et al. Allosteric functioning of dimeric class G-protein-coupled receptors. *FEBS J* 2005; 272:12 2947–2955.
21. Galvez T, Duthey B, Kniazeff J, et al. Allosteric interactions between GB1 and GB2 subunits are required for optimal GABA$_B$ receptor function. *EMBO J* 2001; 20: 2152–2155.
22. Marshall FH, Jones KA, Kaupmann K, Bettler B. GABA$_B$ receptors–the first 7TM heterodimers. *Trends Pharmacol Sci* 1999; 20: 396–399.
23. Gomes I, Gupta A, Filipovska J, et al. A role of heterodimerization of mu and delta opiate receptors in enhancing morphine analgesia. *Proc Natl Acad Sci U S A* 2004; 101: 5135–5139.
24. Rios C, Gomes I, Devi LA. Mu opioid and CB1 cannabinoid receptor interactions: reciprocal inhibition of receptor signaling and neuritogenesis. *Br J Pharmacol* 2006; 148: 387–395.
25. Li X, Staszewski L, Xu H. Human receptors for sweet and umami taste. *Proc Natl Acad Sci U S A* 2002; 99: 4692–4696.
26. Mei B. Dimerization of G-protein-coupled receptors: roles in signal transduction. *Cell Signal* 2004; 16: 175–186.
27. Maggio R, Barbier P, Colelli A, et al. G protein-linked receptors: pharmacological evidence for the formation of heterodimers. *J Pharmacol Exp Ther* 1999; 291: 251–257.
28. Franco R, Ferre S, Agnati L, et al. Evidence for adenosine/dopamine receptor interactions: indications for heteromerization. *Neuropsychopharmacology* 2000; 23(4 suppl): S50–S59.
29. Gomes I, Jordan BA, Gupta A, et al. Heterodimerization of mu and delta opioid receptors: a role in opiate synergy. *J Neurosci* 2000; 20: RC110.
30. Berque-Bestel I, Lezoualc'h F, Jockers R. Bivalent ligands as specific pharmacological tools for G protein-coupled receptor dimers. *Curr Drug Discov Technol* 2008; 5: 312–318.
31. Ballet S, Pietsch M, Abell AD. Multiple ligands in opioid research. *Protein Pept Lett* 2008; 15: 668–682.

32. Gupta A, Décaillot FM, Devi LA. Targeting opioid receptor heterodimers: strategies for screening and drug development. *AAPS J* 2006; 8: E153–E159.
33. Daniels DJ, Lenard NR, Etienne CL, et al. Opioid-induced tolerance and dependence in mice is modulated by the distance between pharmacophores in a bivalent ligand series. *Proc Natl Acad Sci U S A* 2005; 102: 19208–19213.
34. Bhushan RG, Sharma SK, Xie Z, et al. A bivalent ligand (KDN-21) reveals spinal delta and kappa opioid receptors are organized as heterodimers that give rise to delta(1) and kappa(2) phenotypes. Selective targeting of delta-kappa heterodimers. *J Med Chem* 2004; 47: 2969–2972.
35. Xie Z, Bhushan RG, Daniels DJ, Portoghese PS. Interaction of bivalent ligand KDN21 with heterodimeric delta-kappa opioid receptors in human embryonic kidney 293 cells. *Mol Pharmacol* 2005; 68: 1079–1086.
36. Brady AE, Jones CK, Bridges TM, et al. Centrally active allosteric potentiators of the M4 muscarinic acetylcholine receptor reverse amphetamine-induced hyperlocomotor activity in rats. *J Pharmacol Exp Ther* 2008; 327: 941–953.
37. Wang L, Martin B, Brenneman R, et al. Allosteric modulators of G-protein-coupled receptors: future therapeutics for complex physiological disorders. *J Pharmacol Exp Ther* 2009; 331: 340–348.
38. AbdAlla S, Lother H, el Massiery A, Quitterer U. Increased AT(1) receptor heterodimers in preeclampsia mediate enhanced angiotensin ii responsiveness. *Nat Med* 2001; 7: 1003–1009.
39. Barki-Harrington L, Luttrell LM, Rockman HA. Dual inhibition of beta-adrenergic and angiotensin II receptors by a single antagonist: a functional role for receptor-receptor interaction in vivo. *Circulation* 2003; 108: 1611–1618.
40. McGraw DW, Mihlbachler KA, Schwarb MR, et al. Airway smooth muscle prostaglandin-EP1 receptors directly modulate beta2-adrenergic receptors within a unique heterodimeric complex. *J Clin Invest* 2006; 116: 1400–1409.
41. Abul-Husn NS, Sutak M, Milne B, Jhamandas K. Augmentation of spinal morphine analgesia an inhibition of tolerance by low doses of mu-and delta-opioid receptor antagonists. *Br J Pharmacol* 2007; 151: 877–887.
42. González-Maeso J, Ang RL, Yuen T, et al. Identification of a serotonin/glutamate receptor complex implicated in psychosis. *Nature* 2008; 452: 93–97.
43. Schmidt A, MacColl R, Lindau-Shepard B, et al. Hormone-induced conformational change of the purified soluble hormone binding domain of follitropin receptor complexed with single chain follitropin. *J Biol Chem* 2001; 276: 23373–23381.
44. Vignal E, Blangy A, Martin M, et al. Kinectin is a key effector of RhoG microtubule-dependent cellular activity. *Mol Cell Biol* 2001; 21: 8022–8034.
45. Lu J, Chuntharapai A, Beck J, et al. Structure-function study of the extracellular domain of the human IFN-alpha receptor (hIFNAR1) using blocking monoclonal antibodies: the role of domains 1 and 21. *J Immunol* 1998; 160: 1782–1788.
46. Morgan EL, Ember JA, Sanderson SD, et al. Anti-C5a receptor antibodies. Characterization of neutralizing antibodies specific for a peptide, C5aR-(9-29), derived from the predicted amino-terminal sequence of the human C5a receptor. *J Immunol* 1993; 151: 377–388.
47. Lebesgue D, Wallukat G, Mijares A, et al. An agonist-like monoclonal antibody against the human beta2-adrenoceptor. *Eur J Pharmacol* 1998; 348: 123–133.
48. Salle L, Eftekhari P, Aupart M, et al. Inhibitory activity of antibodies against the human atrial 5-HT(4) receptor. *J Mol Cell Cardiol* 2001; 33: 405–417.
49. Blanpain C, Vanderwinden JM, Cihak J, et al. Multiple active states and oligomerization of CCR5 revealed by functional properties of monoclonal antibodies. *Mol Biol Cell* 2002; 13: 723–737.
50. Bozon V, Di Scala E, Eftekhari P, et al. Agonist-like activity of antibodies directed against the second extracellular loop of the human cardiac serotonin 5-HT4(e) receptor in transfected COS-7 cells. *Receptors Channels* 2002; 8: 113–121.
51. Peter JC, Eftekhari P, Billiald P, et al. scFv Single chain antibody variable fragment as inverse agonist of the beta2-adrenergic receptor. *J Biol Chem* 2003; 278(:):36740-36747.
52. Takao M, Iwasa T, Yamamoto H, et al. Anti–bovine rhodopsin monoclonal antibody recognizing light-dependent structural change. *Zoolog Sci* 2002; 19: 651–659.
53. Gupta A. Décaillot F, Gomes I, et al. Conformation state-sensitive antibodies to G-protein-coupled receptors. *J Biol Chem* 2007; 282: 5116–5124.
54. Gupta A, Rozenfeld R, Gomes I, et al. Post-activation-mediated changes in opioid receptors detected by N-terminal antibodies. *J Biol Chem* 2008; 283: 10735–10744.
55. Heimann AS, Gomes I, Dale CS, et al. Hemopressin is an inverse agonist of CB_1 cannabinoid receptors. *Proc Natl Acad Sci U S A* 2007; 104: 20588–20593.
56. Gupta A, Heimann AS, Gomes I, Devi LA. Antibodies against G-protein coupled receptors: novel uses in screening and drug development. *Comb Chem High Throughput Screen* 2008; 11: 463–467.
57. Dragun D, Philippe A, Catar R, Hegner B. Autoimmune mediated G-protein receptor activation in cardiovascular and renal pathologies. *Thromb Haemost* 2009; 101: 643–648.
58. Gomes I. Gupta A, Singh SP, Sharma SK. Monoclonal antibody to the delta opioid receptor acts as an agonist in dual regulation of adenylate cyclase in NG108-15 cells. *FEBS Lett* 1999; 456: 126–130.
59. Mijares A, Lebesgue D, Wallukat G, Hoebeke J. From agonist to antagonist: Fab fragments of an agonist-like monoclonal anti-beta 2-adrenoceptor antibody behave as antagonists. *Mol Pharmacol* 2000; 58: 373–379.
60. Deng HB, Yu Y, Pak Y, et al. Role for the C-terminus in agonist-induced mu opioid receptor phosphorylation and desensitization. *Biochemistry* 2000; 39: 5492–5499.
61. Almagro JC, Fransson J. Humanization of antibodies. *Front Biosci* 2008; 13: 1619–1633.
62. Carter P. Potent antibody therapeutics by design. *Nat Rev Immunol* 2006; 6: 343–357.

Pharmacokinetic/ Pharmacodynamic-Driven Drug Development

James M. Gallo, PhD

Mount Sinai School of Medicine, New York, NY

OUTLINE

- DRUG DEVELOPMENT AND THE GOVERNMENT: A TIME FOR CHANGE
- HISTORICAL ROLE OF PHARMACOKINETICS/PHARMACODYNAMICS IN DRUG DEVELOPMENT
- TOWARD INTEGRATION OF PHARMACOLOGY AND EFFICACY
- THE BIG PICTURE AND THE PHARMACOKINETIC/ PHARMACODYNAMIC BRIDGE

ABSTRACT

The drug discovery and development enterprise, traditionally an industrial juggernaut, has spanned into the academic arena that is partially motivated by the National Institutes of Health Roadmap highlighting translational science and medicine. Because drug discovery and development represents a pipeline of basic to clinical investigations, it meshes well with the "bench to the bedside" prime directive of translational medicine. The renewed interest in drug discovery and development in academia provides an opportunity to rethink the hiearchary of studies with the hope of improving the staid approaches that have been criticized for lacking innovation. One area that has received limited attention concerns the use of pharmacokinetic and pharmacodynamic studies in the drug-development process. Using anticancer drug development as a focus, this review will address past and current deficencies in how pharmacokinetic/pharmacodynamic studies are conducted and offer new strategies that might bridge the gap between preclinical and clinical trials. ***Mt Sinai J Med 77:381–388, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: anticancer drugs, drug development, models, pharmacodynamics, pharmacokinetics.

Drug discovery and development is the essential driving force for drug therapy. This force has been promoted and illustrated in various "pipeline" formats, yet until recently has been largely stagnant in terms of offering innovation to a process that many claim as onerous and inefficient. Even with the renewed attention to the drug-development process, which now has erupted beyond the favored bastion of the pharmaceutical industry into academic centers, there are areas still entrenched in old paradigms. This can be appreciated in the anticancer drug-development area, where efficacy models still dominate and drive the process and a meaningful connection between preclinical and clinical pharmacology is marginal. In this review, the role of pharmacokinetics (PKs) and pharmacodynamics (PDs) in anticancer drug development will be highlighted and a foundation will be cast that places pharmacokinetic/pharmacodynamic (PK/PD) investigations central to the drug-development process.

Address Correspondence to:

James M. Gallo
Department of Pharmacology
and Systems Therapeutics
Mount Sinai School of Medicine
New York, NY
Email: james.gallo@mssm.edu

DRUG DEVELOPMENT AND THE GOVERNMENT: A TIME FOR CHANGE

The drug-development process is intertwined with government regulations and the US Food and Drug Administration (FDA). The FDA has a primary directive to ensure that drugs are safe and effective and to maintain essential standards in clinical trials, yet there is an expectation and growing demand to accelerate

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20193

© 2010 Mount Sinai School of Medicine

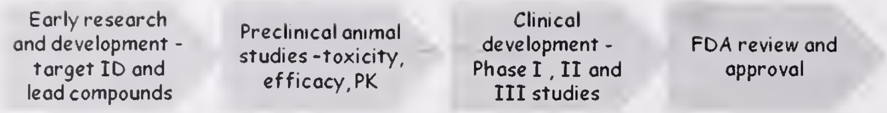

Early research and development - target ID and lead compounds
Preclinical animal studies -toxicity, efficacy, PK
Clinical development - Phase I , II and III studies
FDA review and approval

Fig 1. The drug discovery and development pipeline, showing the major activities of each component. The overall journey for a single drug to gain a new-drug application takes about 10 years and costs hundreds of millions of dollars (adapted from Reference 23). **Abbreviations:** FDA, US Food and Drug Administration; ID, identification; PK, pharmacokinetics.

the drug-development process. If one considers the drug-development pipeline (see Figure 1) and the pharmaceutical industry's mandate to gain regulatory approvals rapidly, one can understand why drug-development science needs to be rethought. The pipeline starting with early drug discovery has been supported by basic science and its fundamental discoveries in biology and "omic" technologies that have afforded a rich milieu for disease targets and new chemical entities. This basic-science enterprise has traditionally been supported by academic institutions in conjunction with National Institutes of Health (NIH) research grants. The data indicate that this paradigm is deficient, with research-and-development expenditures rising and the number of new drug approvals stagnant, and in fact at a low since 2004.[1] The FDA in 2004 initiated the Critical Path Initiative to improve the drug-development process and incorporate the latest scientific advances into the process, and it mostly focused on biomarkers, bioinformatics, and clinical design.[2] In somewhat parallel fashion, the NIH Roadmap was launched as a means to transform biomedical research to ensure basic scientific discovery impacted the treatment of diseases. Thus, both government agencies were trying to grapple with significant advances in scientific discovery and the public demand for tangible benefits, which led to cross-fertilization of research disciplines and an emphasis on translational science. The fallout from these large initiatives did lead to a renewed appreciation of quantitative pharmacology in translational medicine, and terms such as "pharmacometrics" and "model-based drug development" were ever present; yet these procedures found traction in the clinical domain,[3,4] and much of the standard way of doing "business" in preclinical animal testing was unchanged.

HISTORICAL ROLE OF PHARMACOKINETICS/PHARMACODYNAMICS IN DRUG DEVELOPMENT

Most often, PKs and PDs operated independently in the drug-development arena, with PK studies the dominant focus for the purposes of meeting regulatory requirements. This facet is partially responsible for the rather mundane approach to PKs in drug development, as the regulatory "bar" was set low. There was little consideration of how to extend and integrate required PK analyses with the PD properties of a drug candidate. Rather than establishing PK/PD relationships that define efficacy, the FDA mandated studies to provide PK characteristics, referred to as Absorption, Distribution, Metabolism, and Excretion (ADME) of new single agents, issued in the forms of "guidances" that emphasized systemic PK characteristics, toxicokinetics, and safety. Given these guidelines, there was little motivation for companies to examine drug distribution in a systematic manner or in particular to understand drug disposition in target tissues. For anticancer drugs, in which the tumor serves as the target tissue, it is alarming that well-defined protocols were not established to characterize tumor-specific PK properties.

Given these guidelines, there was little motivation for companies to examine drug distribution in a systematic manner or in particular to understand drug disposition in target tissues.

Although the lack of target-tissue PK investigations by pharmaceutical companies may be rationalized from a perspective of meeting regulatory obligations (ie, "get the drug to market"), were academic bodies undertaking such studies? The answer is no, of course, partially explained by the fact that the academic motor for drug discovery and development has only started. A position paper by pediatric oncologists and scientists in 2002 indicated a strong reliance on tumor-efficacy models in their preclinical drug-development scheme.[5] Although the placement of these studies would have been post–Investigational New Drug (post-IND), the scheme, which indicated the use of numerous tumor xenograft and orthotopic models on the road to identifying effective agents to test in the clinic, reflected the prevailing attitude that PK/PD studies were tangential to the process, rather than integral. The proposed PK studies were not extensive and were not designed to build PK models or as a means to link to PD endpoints. Regardless of whether anticancer drug evaluations were pre-IND or post-IND or conducted in industrial or academic centers, the emphasis was to demonstrate tumor shrinkage compared with control. Although information can be gained from tumor size–based efficacy trials, the semiempirical nature of the investigations creates large gaps in our understanding of critical pharmacological components. Drugs deemed ineffective may have failed for PK or PD causes, or both, and without detailed measurements the opportunities to improve and learn from the deficiency, for example, by feedback to medicinal chemists in drug discovery, are lost. Drugs found to significantly reduce tumor growth and prolong animal survival would progress through the drug-development system, yet without a pharmacological foundation. Questions about effective tumor drug concentrations and the extent of target inhibition remained elusive. This PK/PD deficiency compounds the semiempirical approach to drug development in a number of ways. First, because essentially all cancer chemotherapy is a quest to devise effective multidrug combinations, the addition of other drugs to the mix is happenstance, as there are no landmark tumor drug concentrations that should be achieved and no basis to assess drug interactions. Selection of effective doses for compound A and compound B given individually in a combination may be logical, but if the combination fails there is no basis to understand the cause, not to mention how the various permutations of a combination of A + B that entail dose levels, sequences, and duration of therapy might define an optimal regimen. Second, tumor size–based efficacy trials without PK/PD limit the ability to translate preclinical findings to the clinic. An effective mg/kg dose in a mouse cannot be used as the starting dose in a phase I trial of an anticancer agent. Starting doses are based on some fraction of doses causing death, such as a 10% lethal dose (LD10). By establishing internal endpoints, tumor drug concentrations, or target inhibition, extrapolation procedures can be employed to initiate PK/PD-driven clinical trials. The lack of preclinical PK/PD data is not simply an issue in choosing starting doses for patients, but curtails the usefulness of available PK/PD data obtained from patients. It is understandable that measurements of tumor drug concentrations and PD responses in patients will be sparse, singular timed samples that present a challenge as to how to integrate across subjects into meaningful assessments of drug delivery to the tumor and target inhibition. Preclinical PK/PD models that have been derived from tumor measurements provide a means to scale up the models to patients and further allow the limited observations in patients to be beneficial, either to revise the scaled models or to validate them (see below).

To summarize, the major pitfalls in the anticancer drug-development pipeline are minimal focus on drug disposition and dynamics in tumors and a reliance on tumor-size or survival efficacy studies. Even though advances in drug discovery continue, as is evidenced by an abundance of new drug entities and new cellular targets to pursue, there is a breakdown in the in vivo phase that has been plagued by a lack of integration between PK/PD and efficacy studies (see Figure 2). As highlighted in Figure 2, early preclinical drug development is viewed as well integrated and able to produce sufficient lead-candidate compounds to a late preclinical drug development that lacks integration of efficacy and PK studies. The crucial junction of the animal investigations has consequences both in the preclinical and clinical drug-development phases and offers opportunities to enhance its quantitative nature that will permit a translational conduit between early preclinical and clinical investigations.

> *The major pitfalls in the anticancer drug development pipeline are minimal focus on drug disposition and dynamics in tumors and a reliance on tumor-size or survival efficacy studies.*

> *The crucial junction of the animal investigations has consequences both in the preclinical and clinical drug-development phases and offers opportunities to enhance its quantitative nature that will permit a translational conduit between early preclinical and clinical investigations.*

TOWARD INTEGRATION OF PHARMACOLOGY AND EFFICACY

There are efforts to improve drug-development strategies. A primary deterrent of a drug candidate's

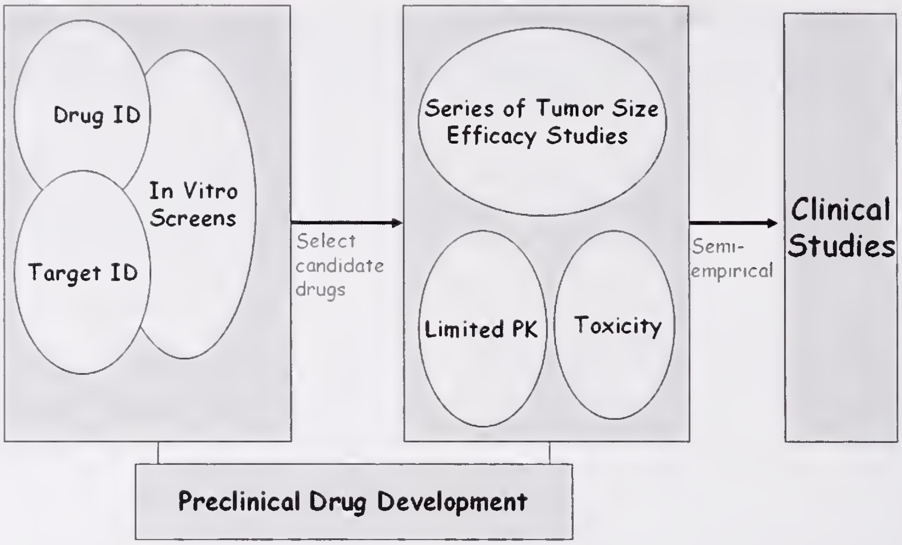


Fig 2. A scheme of traditional anticancer drug development. The early preclinical phase is well-integrated and has applied high-throughput screening tools to the identification of drugs and targets, and in vitro screens to generate viable lead compounds for the animal-study phase. This phase lacks integration and underutilizes PK/PD investigations so that agents entering clinical trials do so in a semiempirical manner. **Abbreviations:** ID, identification; PK, pharmacokinetics; PK/PD, pharmacokinetic/pharmacodynamic.

suitability to progress in drug development is poor oral bioavailability. As early as 1995, Amidon and coworkers[6] devised the Biopharmaceutical Classification System (BCS) for drugs based on their aqueous solubility and membrane permeability, properties determined in vitro, to predict oral absorption. The BCS, which categorized compounds into 4 classes, was adopted by the FDA as a means to assist and bypass certain in vivo bioavailability studies. Wu and Benet[7] proposed the Biopharmaceutical Drug Disposition Classification System as an enhancement to the BCS that accounted for routes of drug elimination and active membrane transport processes afforded by numerous gastrointestinal membrane transporters. These efforts coming from academicians highlight an interesting avenue that is increasingly used in drug-development efforts; namely, the use of in vitro:in vivo correlations. It can be appreciated that the ability to reliably predict in vivo ADME properties from in vitro studies can expedite drug development and save resources.

Academic research has opened an interesting avenue increasingly used in drug development efforts; namely, the use of in vitro:in vivo correlations.

There is considerable interest to employ in vitro methods to predict blood-brain barrier (BBB) permeability for drugs considered for neurotherapeutic areas, such as Alzheimer disease and multiple sclerosis.[8] Although numerous cell systems have been analyzed, including cells of non-neural origin, there is not a single system that has been adopted, in part due to the complex nature of the BBB. In deriving a suitable system for in vitro analyses, recapitulation of the tight endothelial cell junctions and electrical resistance, presence of functioning membrane transporters, and junctional adhesion proteins are significant hurdles. The challenge of an in vitro system to predict BBB permeability is further compromised by the possibility of needing fresh or low-passage endothelial cells. Nonetheless, in vitro models for BBB permeability have been used successfully in certain cases, and future interest is likely to increase as such a tool can be used as feedback to optimize chemical structure for favorable BBB penetration.

Given the growth of pharmacogenetic-PK relationships that are identified in the clinic after drug approval, there has been a proposal to incorporate in vitro screening of drug metabolism and drug transport characteristics of new chemical entities early in preclinical evaluations.[9] This is not a revolutionary concept, as in vitro screens for metabolism and membrane transport are already part of the drug-development machinery. What may be new is the use of a broad panel of in vitro metabolic or transport pathway evaluations that for those found to influence drug disposition can be further queried and tested for polymorphic variations to firmly establish pharmacogenetic-PK relationships early in development. The potential benefits involve not only early identification of problem compounds, but also improved means to evaluate drug interactions and dosing paradigms for desirable agents.

The aforementioned in vitro screening procedures, and others like these, are likely to be beneficial at least in particular therapeutic areas, but these approaches do not address the real problem of meshing in vivo PK and efficacy studies (see Figure 2). Similar to the advances in drug discovery that provide ample new chemical entities, the in vitro screens are improved and may alleviate certain bottlenecks in the drug-development pipeline; yet the primary culprit of poorly integrated preclinical development is animal testing. Before considering alternate in vivo strategies, it should be appreciated that in vitro PD endpoints have not been adequately incorporated into the drug-development hierarchy. This is particularly apparent in the area of in vitro:in vivo correlations, where the effort is focused on in vitro:in vivo relationships of ADME properties. Biomarkers and PDs have received considerable attention in the clinical domain, and there is controversy as to their utility[10,11]; however, what seems to be missing is a concerted effort to translate in vitro cell-based preclinical findings into tumor-based animal investigations. Such an approach was demonstrated for the epidermal growth factor receptor (EGFR) inhibitor gefitinib, whereby phosphorylation of extracellular signal-regulated kinase (pERK) was an essential marker indicative of drug activity, which was further extended to simulate PD profiles in brain-tumor patients.[12,13]

> *Similar to the advances in drug discovery that provide ample new chemical entities, the in vitro screens are improved and may alleviate certain bottlenecks in the drug-development pipeline; yet the primary culprit of poorly integrated preclinical development is animal testing.*

One screening approach that has been used to accelerate drug development in animals is referred to as cassette or N-in-1 dosing.[14,15] The technique involves the administration of low doses of multiple candidate compounds to single animals and monitoring drug concentrations in plasma and sometimes, in addition, tissues. The method was originally employed to assess systemic PK properties of oral bioavailability and total clearance; however, in selected circumstances drug-concentration measurements in tissues have been completed to yield a measure of drug distribution. This latter scenario is most often applied to characterize BBB permeability for neuropharmacological agents.[16] The method has been criticized for the possibility of providing erroneous data due to competition for transport and elimination mechanisms between chemical analogues. There have been examples that show the PK parameters obtained in a cassette format are relatively comparable to those obtained following single drug administrations.[17] Given the low doses and a trend in using smaller-size cassettes, the likelihood of PK drug interactions is reduced and a valid risk, particularly because the method is aptly known as a screening methodology.

The cassette-dosing protocol is a tool that can partially alleviate the bottleneck of in vivo animal testing to provide key PK parameters; however, it does not address the reliance on anticancer drug efficacy studies or allow integration between PK/PD studies and tumor size–based investigations. One approach that we propose is illustrated in Figure 3, tailored to anticancer drugs that may be useful in the treatment of brain tumors. Steps 1–3 represent a cascade of in silico, in vitro, and in vivo studies designed to identify PK/PD failures. The proposed studies overlap with existing screening procedures as generically shown in Figure 2, yet steps 1–3 delineate a progression of investigations that utilize cassette dosing to determine BBB penetration, and a PD phase that seeks to define PD endpoints that respond in a concentration-dependent manner and can also be applied to in vivo tumor models highlighted in step 4. The in silico screen depicted in step 1 utilizes a computer program (ie, ADMET Predictor) that, given the compound's chemical structure, can predict numerous ADME parameters that can assist in the selection of compounds for further investigation. The pitfall with such methods is that the accuracy of the predictions is dependent on the similarity between the molecular descriptors of the test compounds and those compounds used to derive the models internal to the program. The program does flag test compounds that are outside the model's chemical space and had to extrapolate to obtain the final output. The cassette-dosing studies in step 2 can provide the data to build predictive models for BBB permeability for chemotypes of interest. The intended step 3 PD studies are conducted when putative targets have been

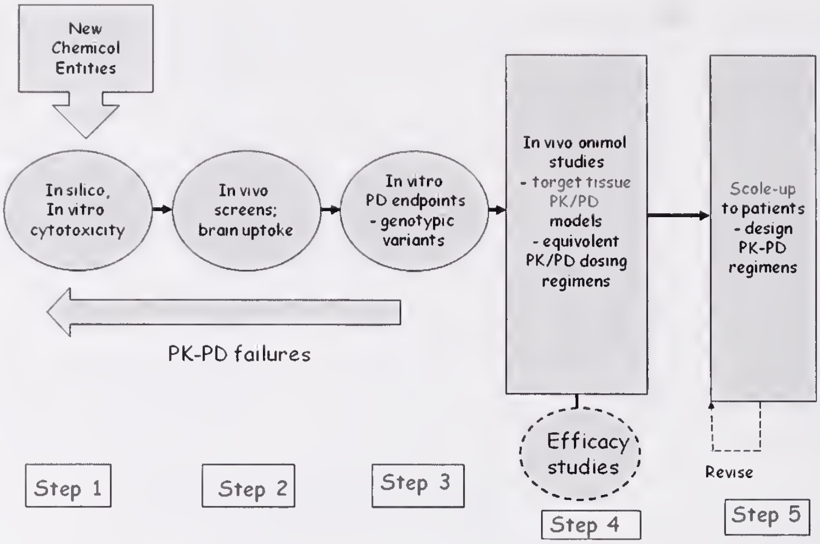

New Chemical Entities
In silico, In vitro cytotoxicity
In vivo screens; brain uptake
In vitro PD endpoints - genotypic variants
In vivo animal studies - target tissue PK/PD models - equivalent PK/PD dosing regimens
Scale-up to patients - design PK-PD regimens
PK-PD failures
Efficacy studies
Revise
Step 1
Step 2
Step 3
Step 4
Step 5

Fig 3. A proposed PK/PD-driven anticancer drug development scheme. Steps 1–3 identify PK/PD failures due to poor systemic properties, brain distribution, or insufficient target inhibition. Step 4 studies are initiated with detailed PK/PD investigations that permit the development of physiologically based PK/PD models that provide a quantitative basis to enter clinical testing. **Abbreviations:** PD, pharmacodynamic; PK/PD, pharmacokinetic/pharmacodynamic.

identified and make use of antibody-based assays in an intermediate or mesoscale format. For targeted tyrosine kinase inhibitors, key signaling molecules in pertinent pathways can be queried in genotyped cells to identify molecular determinants of drug action, which may further serve as PD endpoints in vivo.

The most radical departure from standard drug-development approaches is in step 4, where comprehensive PK/PD studies are completed in orthotopic brain-tumor models in mice prior to any efficacy testing. A protocol design of 2 dose levels both as single dose and multiple-dose regimens in brain tumor–bearing mice would permit the development of hybrid physiologically based (PB) PK/PD models.[12,18] The number and nature of the brain tumor models would be driven by step 3, which could lead to definition of molecular signatures for drug activity or more rudimentary sensitive and resistant cell lines. By distilling the drug concentration and PD measurements into PB-PK/PD models, additional mouse tumor models can be analyzed, possibly employing an equivalent PK/PD dosing strategy that attempts to "normalize" PD responses based on the molecular determinants with an overall goal to establish a robust PK/PD signature for activity.[12] Step 4 also contains a provision to conduct standard efficacy studies; however, these would be driven by the PK/PD models to select dosing regimens to achieve desired PD profiles. Both the extent and duration of inhibition can be simulated via the model. The PD input can be used as input into an efficacy model, most readily done with continuous variables such as tumor size,[13] and ultimately serves as confirmation that the chosen PD endpoints have a meaningful link to accepted gross measures of efficacy.

THE BIG PICTURE AND THE PHARMACOKINETIC/ PHARMACODYNAMIC BRIDGE

The drug-development pipeline fails as a translational tool unless there is a mechanism to apply the information gathered in the preclinical stages to the clinic. The standard approach (Figure 2), discussed above, fails and at best is a semiempirical leap from the lab to the clinic. The alternate approach (Figure 3), aptly referred to as a PK/PD-driven drug-development strategy, offers a quantitative foundation to transition from the lab to the clinic. The PB-PK/PD models can be built in a global or hybrid format depending on the extent of tissue data obtained, but in either case the target tissue

The drug-development pipeline fails as a translational tool unless there is a mechanism to apply the information gathered in the preclinical stages to the clinic.

is represented physiologically with organ blood volume and volumes,[19] and thus minimizes the "black box" compartmental modeling approach that prevents interspecies scale-up. Normal human and preferably patient values replace the corresponding physiological parameters in the animal-based PB-PK/PD models. The challenge in the scale-up procedure is how to estimate the drug-related parameters including clearance, drug distribution, and transport variables. There is no definitive treatise on how this can be accomplished, but parameters determined from drug-distribution studies in animals are clearly applicable.[18] Consistent with the original ideas of PB-PK modeling and the recent resurgence in the approach in the drug-development domain, a combination of in silico and in vitro methods may provide reasonable approximations for drug-dependent parameters that may minimize animal testing or relegate them to validation studies.[20] What has not been rigorously addressed is both the need and method to scale up parameters associated with the PD model. This is a tantalizing subject, and one might argue that for certain drug targets only minor adjustments may be needed in patients, because for most targeted drugs, such as tyrosine kinase inhibitors, PD endpoints are represented in the model as a fraction of the baseline rather than in absolute units. The net result of the scale-up procedure is to define a PB-PK or PB-PK/PD model in patients that can simulate drug concentrations in tissues and/or PD response in the tumor as a function of time under any desired conditions.[13,18] The simulation results can be used to set dosing regimens in early clinical trials, and as more patient data becomes available the model can be revised, thus offering a tool that can continually assess the pharmacological aspects of therapy.

We have used a preclinical PB-PK/PD model for gefitinib as a basis to derive an analogous model in brain-tumor patients.[13] In this case, there was sufficient PK data in patients to derive a simple blood flow–limited PB-PK model that was subsequently linked to the preclinical PD model that used phosphorylation of extracellular signal-regulated kinase (pERK) as the endpoint in tumors. The study was constructed to contrast gefitinib multiple-dose regimens that would be required in patients to attain the same inhibition profile of pERK in tumors with wild-type epidermal growth factor receptor (EGFR) and mutant vIIIEGFR, which is the more sensitive tumor.[13] Figure 4 illustrates the results of the modeling procedure, which indicates a 2-fold greater daily dose of gefitinib (1000 mg/day) is needed in patients with wild-type EGFR tumors compared with those with vIIIEGFR tumors. Aforementioned questions on how preclinical PD models may be extrapolated to patients, and how more patient data can be collected in tumors, are areas that warrant future investigations, yet this simulation highlights how a tumor-based PK/PD model could be used to tailor therapy to patients.

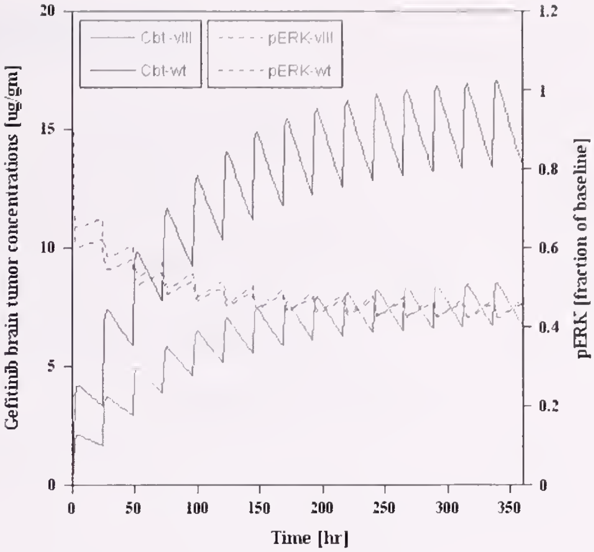


Fig 4. Physiologically based PK/PD model–predicted gefitinib brain tumor concentrations (left y-axis) and fractional inhibition of pERK (right y-axis) in patients with either wild-type EGFR or mutant vIIIEGFR tumors. To obtain an analogous pERK profile in both tumor types, a regimen of 1000 mg/day or 500 mg/day of gefitinib is predicted in the wild-type EGFR and vIIIEGFR groups, respectively. The simulation used a 15-d duration of gefitinib administration (adapted from Reference 13). **Abbreviations:** EGFR, epidermal growth factor receptor; pERK, phosphorylation of extracellular signal-regulated kinase; PK/PD, pharmacokinetic/pharmacodynamic; wt, wild type.

The PK/PD-driven approach to drug development is quite consistent with the recent emphasis on individualized therapy,[21,22] as there is a merging of patient covariate and pharmacogenetic information with PK/PD parameters. This area, known as population-based PK, has been embraced by many clinical pharmacologists and the FDA. The fundamental differences in the proposed PK/PD-driven

The PK/PD-driven approach to drug development is quite consistent with the recent emphasis on individualized therapy, as there is a merging of patient covariate and pharmacogenetic information with PK/PD parameters.

drug development scheme and the prevailing individualized therapy via population-based methods are 2-fold: (1) the use of tumor-based PK/PD data, and (2) initiation of the model-building exercise early in preclinical development. In conclusion, derivation of PB-PK/PD models as a key preclinical goal may limit tedious tumor size–based efficacy studies, and as tumor-based models may provide a compelling resource to individualize patient therapy in a rational manner.

ACKNOWLEDGMENTS

The author is supported in part by NIH grants CA127963 and CA127963-03S1.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Kaitin KI. Deconstructing the drug development process: the new face of innovation. *Clin Pharmacol Ther* 2010; 87: 356–361.
2. Woodcock J, Woosley R. The FDA critical path initiative and its influence on new drug development. *Annu Rev Med* 2008; 59: 1–12.
3. Powell JR, Gobburu JV. Pharmacometrics at FDA: evolution and impact on decisions. *Clin Pharmacol Ther* 2007; 82: 97–102.
4. Danhof M, de Lange EC, Della Pasqua OE, et al. Mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling in translational drug research. *Trends Pharmacol Sci* 2008; 29: 186–191.
5. Houghton PJ, Adamson PC, Blaney S, et al. Testing of new agents in childhood cancer preclinical models: meeting summary. *Clin Cancer Res* 2002; 8: 3646–3657.
6. Amidon GL, Lennerns̈ H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharm Res* 1995; 12: 413–420.
7. Wu CY, Benet LZ. Predicting drug disposition via application of BCS: transport/absorption/elimination interplay and development of a biopharmaceutics drug disposition classification system. *Pharm Res* 2005; 22: 11–23.
8. Cecchelli R, Berezowski V, Lundquist S, et al. Modelling of the blood-brain barrier in drug discovery and development. *Nat Rev Drug Discov* 2007; 6: 650–661.
9. Katz DA, Murray B, Bhathena A, Sahelijo L. Defining drug disposition determinants: a pharmacogenetic-pharmacokinetic strategy. *Nat Rev Drug Discov* 2008; 7: 293–305.
10. Glassman RH, Ratain MJ. Biomarkers in early cancer drug development: limited utility. *Clin Pharmacol Ther* 2009; 85: 134–135.
11. Carden CP, Banerji U, Kaye SB, et al. From darkness to light with biomarkers in early clinical trials of cancer drugs. *Clin Pharmacol Ther* 2009; 85: 131–133.
12. Wang S, Guo P, Wang X, et al. Preclinical pharmacokinetic/pharmacodynamic models of gefitinib and the design of equivalent dosing regimens in EGFR wild-type and mutant tumor models. *Mol Cancer Ther* 2008; 7: 407–417.
13. Wang S, Zhou Q, Gallo JM. Demonstration of the equivalent pharmacokinetic/pharmacodynamic dosing strategy in a multiple-dose study of gefitinib. *Mol Cancer Ther* 2009; 8: 1438–1447.
14. White RE, Manitpisitkul P. Pharmacokinetic theory of cassette dosing in drug discovery screening. *Drug Metab Dispos* 2001; 29: 957–966.
15. Smith NF, Raynaud FI, Workman P. The application of cassette dosing for pharmacokinetic screening in small-molecule cancer drug discovery. *Mol Cancer Ther* 2007; 6: 428–440.
16. Zhang MY, Kerns E, McConnell O, et al. Brain and plasma exposure profiling in early drug discovery using cassette administration and fast liquid chromatography-tandem mass spectrometry. *J Pharm Biomed Anal* 2004; 34: 359–368.
17. He K, Qian M, Wong H, et al. N-in-1 dosing pharmacokinetics in drug discovery: experience, theoretical and practical considerations. *J Pharm Sci* 2008; 97: 2568–2580.
18. Zhou Q, Guo P, Kruh GD, et al. Predicting human tumor drug concentrations from a preclinical pharmacokinetic model of temozolomide brain disposition. *Clin Cancer Res* 2007; 13: 4271–4279.
19. Gallo JM, Vicini P, Orlansky A, et al. Pharmacokinetic model-predicted anticancer drug concentrations in human tumors. *Clin Cancer Res* 2004; 10: 8048–8058.
20. De Buck SS, Mackie CE. Physiologically based approaches towards the prediction of pharmacokinetics: in vitro-in vivo extrapolation. *Expert Opin Drug Metab Toxicol* 2007; 3: 865–878.
21. Gerretsen P, Müller DJ, Tiwari A, et al. The intersection of pharmacology, imaging, and genetics in the development of personalized medicine. *Dialogues Clin Neurosci* 2009; 11: 363–376.
22. Ma BB, Hui EP, Mok TS. Population-based differences in treatment outcome following anticancer drug therapies. *Lancet Oncol* 2010; 11: 75–84.
23. O'Driscoll C. A virtual space odyssey. In: Horizon Symposium, *Charting Chemical Space: Finding New Tools to Explore Biology*. http://www.nature.com/horizon/chemicalspace/background/odyssey.html. New York, NY: Nature Publishing Group; 2004; 1–4.

SPECIAL FEATURE

That I Won't Translate! Experiences of a Family Medical Interpreter in a Multicultural Environment

Rhona D. Seidelman, PhD, and Yaacov G. Bachner, PhD

Ben-Gurion University of the Negev, Beer-Sheva, Israel

OUTLINE

Challenges to the Medical Interpreter in a Multicultural Environment
Communication and Trust
Family-Member Interpreters: Drawbacks of Dependence
Professional Medical Interpreters: the Preferable Alternative
Medical Interpreters in Israel: the Bar-Ilan Initiative
a Case in Point: How it Ended

ABSTRACT

Family members used as patients' interpreters are a common occurrence in the medical environments of multicultural societies. It is recognized that the use of the family-member interpreter may have some benefits. However, studies show that this option also has substantial disadvantages and therefore suggest that the use of professional medical interpreters is the preferable option for effective quality care. The purpose of the current study is to present the narrative of a family-member interpreter in Israel, a diverse immigrant society. While numerous studies have been done on the challenges in the doctor-interpreter-patient medical encounter, these studies tend to focus on the experiences of the physicians or the patients, and the perspective of the interpreter is often sidelined. After discussing the various interpreting options, we suggest that the perspective of the family-member interpreter strengthens assertions that professional interpretation is the best option for multilingual medical environments. ***Mt Sinai J Med 77:389–393, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: communication, interpreters, Israel, multiculturalism.

Address Correspondence to:

Rhona D. Seidelman
The Moshe Prywes Center for Medical Education
Faculty of Health Sciences
Ben-Gurion University of the Negev
Beer-Sheva, Israel
Email: seidelma@bgu.ac.il

We begin this paper with an experience of one of the coauthors, Rhona Seidelman. Eight months after Dr. Seidelman's parents immigrated to Israel, her father was admitted to Soroka University Medical Center in Beer-Sheva. Throughout the night, Dr. Seidelman filled the dual role of supportive child and Hebrew-English interpreter. Here, in Dr. Seidelman's words, is how the case began:

"One morning in late March my father called me at home to request that I accompany him and my mother to the emergency room, where he was going to be admitted and would need my help understanding the Hebrew. There was no chance that I would have refused his request. Since having immigrated to Israel from Canada a few months earlier, my parents had been adamant about their independence and were horrified about the possibility of becoming a burden upon their children. So I knew that if they were actually turning to me for help it was because there was no other option.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20189

© 2010 Mount Sinai School of Medicine

"When Dr. Abu [name has been changed] examined my father, we had been in the hospital for nearly 6 hours. By that point I was trying to appear upbeat for my parents, but, like them, I was tired and afraid (there was reason to believe that my father's condition was very serious). To make matters worse, it had become clear that the department was so overcrowded that my father (a very private, reserved Canadian) was going to have to sleep in the hallway. As far as Dr. Abu was concerned, I can only assume that he himself was tired and under pressure, since he was the attending physician for that overflowing department. Dr. Abu was about the sixth hospital employee that my father had spoken to that day (including administrative assistants, nurses, orderlies, and physicians) and, as such, he was the sixth person I had translated for in that 1 day.

"My first impression of the doctor was positive. I was relieved that he would ask me to translate into Hebrew what my father had said–even though he clearly had some command of English. Throughout the day we had encountered several doctors and nurses who refused to use me as interpreter, proudly insisting–in Hebrew–that their English was excellent (which mostly proved not to be the case).

"After explaining to Dr. Abu that he was also a physician, my father then explained, in English, the many complicated steps that had led him there that evening. I translated between them, switching constantly from Hebrew to English and back again. As the examination progressed and the severity of my father's condition became apparent, I became increasingly uncomfortable with my role. At some point it dawned on me how different this was from translating for my father in a bank or at a restaurant; here, an imprecise translation on my part could have critical implications.

"The situation became increasingly complicated when Dr. Abu and my parents became impatient with one another. I don't recall what the trigger was (on a night with so many anxieties and uncertainties, it could have been any number of things) and I don't recall exactly what it was that Dr. Abu said, but my immediate response to him was, "Okay, *that* I won't translate." He had said something about there being too many questions, or something about him needing to go, or something about him already having answered their question. Knowing my parents, I knew that response of his would have broken through the appearance of calm they were barely maintaining, and as their daughter, censorship of the doctor seemed to be my only option.

"Naturally, many other factors beyond language came into play that night. But for my part I had come to fear the role of medical interpreter. The question was, what were my parents' other options? And how do other multicultural, multilingual societies deal with similar problems of doctor-patient communication?"

CHALLENGES TO THE MEDICAL INTERPRETER IN A MULTICULTURAL ENVIRONMENT

In this article, we will use Dr. Seidelman's story as a springboard for exploring the topic of doctor-patient communication via an interpreter; in particular, we will examine the challenges experienced by and posed by family-member interpreters, who are very common players in multilingual medical settings.

In addition to considering the challenges involved in communicating through a medical interpreter, we will also look at various initiatives being undertaken to address these challenges. We would like to suggest that the perspective of the family-member interpreter gives further support to the claims that professional interpretation is the best option for multilingual medical environments. Furthermore, in looking at the local context, we assert that Israel, a vastly diverse immigrant society with a top-level healthcare system, must no longer rely on amateur interpreters in the medical setting. This reliance places unfair responsibility on the interpreters themselves and compromises the quality of care. Initiatives that are beginning to take root in Israel in the training of professional medical interpreters must be nourished and the place of these professionals in the hospital setting must be acknowledged as an incontestable necessity for the provision of proper care.

COMMUNICATION AND TRUST

Doctor-patient communication constitutes the foundation on which proper and effective medical care stands. It is through communication that physicians reach accurate diagnoses, assign appropriate treatment programs, as well as calm and guide those in their care. Patients' trust in physicians is a fundamental element of this relationship.[1] Effective communication enhances patients' trust in their physicians–a factor which Peabody[2] has shown to be crucial in determining the effectiveness of medical care. Studies have shown that patients' lack of a sense of trust will contribute to a wide range of adverse effects, including slow patient recovery, lower satisfaction with the physician and the service, and higher costs to the medical system due to unnecessary tests and recurrent visits.[3,4] Language barriers can impede the creation of a mutually trusting relationship between physicians and patients. The facilitation of doctor-patient communication as a necessary component of quality, cost-effective healthcare is a challenge faced by multicultural societies the world over.

> *Language barriers can impede the creation of a mutually trusting relationship between physicians and patients. The facilitation of doctor-patient communication as a necessary component of quality, cost-effective healthcare is a challenge faced by multicultural societies the world over.*

FAMILY-MEMBER INTERPRETERS: DRAWBACKS OF DEPENDENCE

Various options exist for bridging language barriers in the medical setting, including bilingual health workers, trained interpreters, and untrained volunteers.[5] In addition, relying on family members to act as medical interpreters is a well-known occurrence throughout the world, with advantages of familiarity, comfort, and accessibility.[5–7] Moreover, the hospital benefits by receiving this service for free. Yet, clearly there are also substantial drawbacks to this arrangement. As seen in my comment "*that* I won't translate," there is a realistic possibility that the interpreter will interfere in the doctor-patient exchange. Indeed, studies show that my censorship was a response common among family-member interpreters:

- "the risk is that someone close to the patient will not stop their own views of the situation coloring their translation. For the best of motives they may try to protect the patient from bad news, or decide to tell them later in private [. . .] Occasionally more sinister reasons may be behind altering what is said."[5]
- "Family interpreters acted mostly as participants, often speaking as themselves rather than providing accurate renditions of both the doctor's and patient's comments."[7]

Other studies emphasize the human and financial costs that can arise out of faulty, untrained medical interpretation: "The subtleties of interpretation can be lost by the less skilled and the patient can be misrepresented and then misdiagnosed."[6] Indeed, flawed communication can result in misdiagnosis, mistreatment and deteriorated health.

Another drawback that is less documented in existing literature on the use of untrained interpreters is the harm caused to the actual interpreter. It places on that individual tremendous responsibility that he or she has not been trained to shoulder. One can deduce that the only reason why high-tech hospitals would allow such a scenario to exist is because of the low-tech nature of language. In a hospital setting, I would not be allowed to conduct a CT scan–for the logical reason that I am untrained, and harm could come of giving an untrained individual that power and responsibility. Why then, as an untrained interpreter, should I have been given the responsibility for one of the most vital of diagnostic tools: language?

> *There are substantial drawbacks to the use of family-member interpreters. The interpreters may interfere with the doctor-patient exchange, resulting in misdiagnosis, mistreatment, and deteriorated health. They may also suffer harm themselves.*

PROFESSIONAL MEDICAL INTERPRETERS: THE PREFERABLE ALTERNATIVE

Scientific literature shows that each existing interpretation solution has particular benefits and drawbacks. Nevertheless, there is wide-scale acknowledgement that the use of professional medical interpreters is the most effective option, resulting in improved care, greater patient satisfaction, and lower expenses for the medical system.[8–10] The 2 central needs that professional medical interpreters fulfill are neutrality and accuracy. Professional medical interpreters are the most reliable agent for conveying the information between patient and physician with the most precision and with the least interference. Trained interpreters ensure a skill and particular vocabulary that ad hoc interpreters are not likely to possess.[8–10] Moreover, physicians and patients have reported a greater level of comfort in working with a professional medical interpreter in terms of doctor-patient confidentiality; physicians feel more comfortable addressing issues of sexuality and domestic abuse when a family member is not interpreting, and patients tend to feel more at ease disclosing personal information to a professional, neutral agent rather than to a family member.[9]

The potential drawbacks to working with a professional medical interpreter generally stem from the natural challenges involved in the triad encounter (patient-interpreter-physician), the awkwardness that comes of a third presence in a medical examination, and the frustration of an impeded patient-physician verbal exchange. However, active steps are being taken to address and rectify these challenges through the proposal of practical solutions.[8,11] Certainly professional medical interpretation is not without its flaws, but of all existing options it is the one consistently referred to as superior.[8,10,11]

> *The use of professional medical interpreters results in improved care, greater patient satisfaction, and lower expenses for the medical system. At the same time, the presence of a professional medical interpreter during a physician-patient exchange can create an awkwardness that impedes the exchange.*

MEDICAL INTERPRETERS IN ISRAEL: THE BAR-ILAN INITIATIVE

The cultural diversity of the Israeli population was largely determined by a massive wave of immigration that followed the establishment of the state in 1948. From 1948 to 1961, more than 1 million people migrated to Israel from more than 50 different countries, with the largest groups coming from Iraq, Romania, Morocco, and Yemen.[12] Further influencing the country as it exists today were 2 other significant waves of immigration in the latter part of the 20th century, from Ethiopia and the former USSR. In Israel there are approximately 72,000 Ethiopian-born citizens, while Israelis born in the former USSR make up 10% of the general population, numbering 705,100.[13] Arab citizens of Israel make up 20% of Israel's population. In addition, there are approximately 180,000 migrant workers from more than 100 countries who reside in Israel.[14]

Thus, more than 30% of the general population does not have Hebrew as a first language. And though according to Israeli law a public institution is obligated to supply an interpreter to an individual unable to communicate in Hebrew, this law is not always put into practice. Furthermore, the healthcare system itself is not obligated to offer the services of an interpreter. As a result, language obstacles to doctor-patient communication are widespread in Israel.[4,15]

A program in walk-in clinics in the northern city of Haifa, in which nursing students from the Ethiopian community have been employed as Hebrew-Amharic interpreters, has met with great success.[14] Nevertheless, the situation remains that Israeli hospitals rely primarily upon ad hoc interpreters. Thus, in this hugely diverse immigrant society, the solutions currently available for cases like my father's are reliance on family members and friends, or a bilingual hospital employee, if there is one available who can be taken away from his or her paid job. As discussed earlier, both of these scenarios are severely flawed and may seriously compromise the standards of care.

The solution lies in the recruitment of trained medical interpreters as an integral part of the team of healthcare professionals, a system that can already be found (in various manifestations and to different extents) in other immigrant societies such as the United States, Canada, England, Australia, and Scandinavian countries.

> *The solution lies in the recruitment of trained medical interpreters as an integral part of the team of healthcare professionals, a system currently employed in such immigrant societies as the United States, Canada, England, Australia, and Scandinavian countries and adaptable to others such as Israel.*

Although the medical-interpretation system has its own challenges that must be addressed and improved upon, the success of various international programs and, more locally, the Haifa program, provides successful models to build upon.[7,16,17] Several researchers from Bar-Ilan University's Department of Translation Studies have been leading initiatives to reform Israel's healthcare policy regarding interpreters.[14,15] It is the responsibility of the Israeli healthcare system to actively push the medical interpreters' program out of its infancy and integrate medical interpreters as central, accessible players on hospital wards whose only job is to facilitate communication. It should be unthinkable that communication–such a fundamental aspect of medical care–would be entrusted to nonprofessionals.

A CASE IN POINT: HOW IT ENDED

Here, again in Dr. Seidelman's words, is how her father's story ended:

"Fortunately, my father's story ended well, but we all came away from it very shaken. And I can't help but wonder how our encounter could have been different had he been entitled to the services of a trained medical interpreter. To begin with, my parents could have been put at ease through the knowledge that they were in the hands of a health-care professional rather than an amateur linguist; in this way trust would likely have been fortified and care improved. Additionally, their sense of independence would have been augmented, for were they to know that the hospital had an effective program of medical interpreters that they could rely on at any time, they would gain the option of not even involving their daughter if they did not want to. For my part, a program of medical interpreters could reduce my anxiety and take out of my hands any responsibility for my father's health–responsibility that a hospital should in fact not be giving to untrained individuals. This would allow me to retire from the role of impartial, scientifically minded medical interpreter–a job for which I have no training and to which I am not suited–and I could then focus on being a biased, concerned daughter, a job that suits me so much more."

REFERENCES

1. Thom DH. Training physicians to increase patient trust. *J Eval Clin Pract* 2000; 6: 245–253.
2. Peabody FW. The care of the patient. *JAMA* 1927; 88: 877–882.
3. Bachner YG, Carmel S, Lubetzky H, et al. Therapists' communication styles and parents' global trust in the therapists: a comparison between Jewish and Bedouin parents. *Pediatr Phys Ther.* 2005; 17: 1–7.
4. Bachner YG, Carmel S, Lubetzky H, et al. Parent' therapist communication and satisfaction with the services of a child development center: a comparison between Israeli parents' Jews and Bedouins. *Health Commun* 2006; 19: 221–229.
5. Phelan M, Parkman S. How to work with an interpreter. *BMJ* 1995; 311: 555–557.
6. Jackson C. Medical interpretation: an essential clinical service for non-English-speaking immigrants. In: Loue S, ed. *Handbook of Immigrant Health*. New York, NY: Plenum Press; 1998; 61–79.
7. Schapira L, Vargas E, Hidalgo R, et al. Lost in translation: integrating medical interpreters into the multidisciplinary team. *Oncologist* 2008; 13: 586–592.
8. Chan Y, Alagappan K, Rella J, et al. Interpreter services in emergency medicine. *J Emerg Med* 2010; 38: 133–139.
9. Rosenberg E, Leanza Y, Seller R. Doctor' patient communication in primary care with an interpreter: physician perceptions of professional and family interpreters. *Patient Educ Couns* 2007; 67: 286–292.
10. Rosenberg E, Seller R, Leanza Y. Through interpreters' eyes: comparing roles of professional and family interpreters. *Patient Educ Couns* 2008; 70: 87–93.
11. Fatahi N, Hellström M, Skott C, Mattsson B. General practitioners' views on consultations with interpreters: a triad situation with complex issues. *Scand J Prim Health Care* 2008; 26: 40–45.
12. Wallach J, Lissak M, eds. *Carta's Atlas of Israel, The First Years 1948–1961*. Jerusalem, Israel: Carta; 1978.
13. Israel Central Bureau of Statistics. Annual Statistical Abstract of Israel, 2007. http://www.cbs.gov.il.
14. Schuster M. Linguistic and cultural mediation in public institutions [in Hebrew]. *Hed ha ulpan hadash*. 2007; 91: 1–8.
15. Ravnovsky Y. Where does it *halit* you?: communication difficulties between doctors and patients in the Israeli health care system [in Hebrew]. *Hed ha ulpan ha hadash* 2006; 89: 1–15.
16. Chowdhury U. Listening to the patient. *BMJ* 2000; 320: 1445.
17. Pérez-Stable EJ, Nápoles-Springer A. Interpreters and communication in the clinical encounter. *Am J Med* 2000; 108: 509–510.

ETHICS

Joseph's Wishes: Ethical Decision-Making in Duchenne Muscular Dystrophy

Leigh Penner, LCSW,[1] Rabbi Mollie Cantor, BCC,[1] and Linda Siegel, MD[2]

[1]Mount Sinai Kravis Children's Hospital, New York, NY,
[2]Department of Pediatrics, Albert Einstein College of Medicine, Bronx, NY

OUTLINE

Case Study
Discussion
Disclosures

ABSTRACT

The death of a child is an uncommon occurrence and can be difficult for families to accept. Attempts by the healthcare team to discuss end-of-life issues can be so upsetting to families that they refuse to participate in the discussion. Even with diseases where an early death can be anticipated, such as Duchenne muscular dystrophy, the family is often reluctant to discuss end-of-life issues when the patient is relatively healthy, preferring to focus on the current health issues. In this article we discuss what happens when there are disagreements between the patient and family and the needs of both must be taken into consideration. ***Mt Sinai J Med 77:394–397, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: decision-making, Duchenne muscular dystrophy, ethics, palliative care, pediatric.

Duchenne muscular dystrophy (DMD) is a genetic X-linked recessive degenerative disease primarily affecting boys that is diagnosed in early childhood. Survival is rare beyond the early 30s, as DMD ultimately affects the cardiac and respiratory muscles (see Muscular Dystrophy Association Web site, Duchenne muscular dystrophy page, www.mda.org/disease/dmd.html). In this article we will review a case taken from the Pediatric Advanced Care and Comfort Team (PACC Team) at the Mount Sinai Kravis Children's Hospital. The purpose of this article is to discuss the ethical problems in dealing with decision-making regarding end-of-life care when there is disagreement between the patient and family.

Address Correspondence to:

Leigh Penner
Mount Sinai Kravis Children's Hospital
New York, NY
Email: leigh.penner@mountsinai.org

The PACC Team at Mount Sinai provides palliative-care services to pediatric patients who have potentially life-limiting illnesses and their families. Patients are referred to the PACC Team either through their primary healthcare team or from the pediatric intensive care unit (PICU). The interdisciplinary members of this team are an attending physician, nurse practitioner, social worker, pediatric chaplain, and child life specialist. The team's goal is to offer comprehensive comfort care though the relief of suffering by addressing the physical, emotional, spiritual, and psychosocial needs of patients in their environment. This process often begins with a discussion of goals of care and facilitation of communication between the patient, family, and healthcare professionals.

When a child is sick, the patient and family must be cared for as a single unit. Because the death of a child is unusual and very difficult for parents to accept, attempts by the healthcare team to bring up end-of-life issues can be very upsetting to the parents, who may refuse to participate in the discussion. When doctors ask parents about withdrawing life-sustaining treatment for a child, they are asking the unthinkable. To parents, it may feel as if they are being asked to "kill their child." Even when an early death can be anticipated, as in a patient with

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20196

© 2010 Mount Sinai School of Medicine

When doctors ask parents about withdrawing life-sustaining treatment for a child, they are asking the unthinkable. To parents, it may feel as if they are being asked to "kill their child."

DMD, the family is often reluctant to discuss end-of-life issues when the patient is relatively healthy, preferring to focus on the current health issues.

CASE STUDY

Joseph was a 21-year-old, single male diagnosed at age 3 with DMD. He lived in an urban home with his single mother and 2 younger brothers. Joseph enjoyed being at home with his family. According to his mother, he was seen as a family leader and role model. He was the eldest male, a high school graduate, and provided a focal point for how his family organized their lives. Although his body had become compromised, his mental functioning remained intact and he was able to communicate with friends and family. Joseph's family believed in God and considered themselves Christians. His mother identified her faith as a resource for strength and coping.

At the time of diagnosis, the life-limiting prognosis for DMD was explained to Joseph's mother. Due to his young age, Joseph was not involved in any healthcare-related discussions. Over the course of many years and consistent with the clinical trajectory of DMD, Joseph's muscles became progressively weaker and his heart function decreased significantly. He was primarily homebound, cachectic, contracted, and unable to walk. He required assistance with all aspects of self-care. Despite these limitations, he continued to be engaged with his younger brothers and extended family members. He enjoyed playing video games, watching television, and joking around with his brothers.

During his illness, Joseph was followed at Mount Sinai Hospital's Outpatient Pediatric Cardiology Clinic and had multiple admissions to Mount Sinai's PICU. During these hospitalizations, the PACC Team tried to engage Joseph's mother in making end-of-life decisions. She would not participate in these conversations, and she could not tolerate any discussion that highlighted the fact that her child would die.

On his final admission, Joseph's mother was still unwilling to make any end-of-life decisions. Joseph presented in respiratory failure and was emergently intubated in the emergency room. Joseph did not have an advance directive delineating his wishes with respect to intubation, mechanical ventilation, and possible tracheostomy for long-term ventilation. His mother refused to make a decision as to which interventions should be introduced or withdrawn, and she would not offer any opinions as to what the next steps in Joseph's care should be. She was paralyzed by the enormity of the decisions and appeared content to maintain the status quo with Joseph intubated and on the ventilator.

Joseph, the patient in this case, was an adult and had the right to participate in decisions regarding his healthcare.

Joseph was an adult and had the right to participate in decisions regarding his healthcare. He was well known to the psychosocial team, which on previous admissions had assessed him as having the ability to make independent decisions. He had the right to decide what medical interventions he wanted and how he wanted to be treated as his disease progressed. Joseph's mother was aware that the team would meet with Joseph, and she did not resist this intervention.

The attending physician, social worker, child life specialist, and chaplain were all present for the initial conversation with Joseph about his medical treatment. He was unable to speak or write, but used simple sign language and head movements to communicate his thoughts. Joseph's mother chose not to be present. Joseph was able to affirm that he understood the medical terminology being used, his diagnosis, and his prognosis. He communicated that he had known someone with a tracheostomy and had known others with DMD who had died. The team explained that if he did have a tracheostomy, he would also need a ventilator and would potentially need to enter a chronic-care facility. When asked if he wanted a tracheostomy, he answered no. When asked if he wanted to be extubated, he answered yes. Joseph also expressed that he wanted to be extubated immediately. When asked if he understood that without the support of the ventilator he would die, he affirmed that he understood. With Joseph's permission, the team met with his mother to share with her Joseph's wishes. Joseph's mother was comfortable letting Joseph make this decision and she accepted his decision, albeit with much sadness. In the end she said it was "all left up to God." This statement seemed to allow her to maintain her faith and simultaneously avoid active participation in the decision-making process.

A conversation was then facilitated between Joseph and his mother. It was clear to all of us how painful and difficult this was for her. However, she was able to hear what Joseph wanted and agreed to honor his wishes. At this point, Joseph's mother requested time to speak with her other children and extended family members.

Joseph's youngest brother shared with us how much Joseph had enjoyed being outside. Given his physical state, Joseph had been unable to leave his home for some time. We were able to take Joseph and his brother to our rooftop garden so he could enjoy the sun again. It was a special time not only for Joseph and his brother, but also for all of us who were caring for him. It was an opportunity for Joseph to say goodbye to the world.

The day before Joseph died, he was extubated while surrounded by his family. He spent the next hours enjoying their company, watching TV, and joking with his brothers. During this time he was asked again if he wanted to be reintubated or have a tracheostomy. Again he answered no. We asked again whether he understood that it meant he would die. Again, he answered yes. On the second day of his extubation, Joseph spent the morning with his family. Shortly thereafter, Joseph had a peaceful death. The team believed that he did not suffer from pain or anxiety.

DISCUSSION

Deciding to withdraw a life-sustaining treatment is not an easy decision to make. Even the language we use when discussing withdrawal can complicate how one views it. In this case, was the ventilator providing life-sustaining treatment or prolonging the dying process?

Was the ventilator in this case providing life-sustaining treatment or prolonging the dying process?

Would removing the ventilator hasten Joseph's death or allow a natural death? A competent person has an established right to refuse life-saving treatment. The goal of the healthcare team was to lessen Joseph's suffering and honor his desires concerning his medical care. After ascertaining Joseph's wishes, we did not want to prolong his suffering and not respect the choice he had made. He had a legal right to refuse treatment, and his caretakers had an ethical duty to respect his autonomy.

Joseph had a legal right to refuse treatment, and his caretakers had an ethical duty to respect his autonomy.

He was awake and alert, and he had expressed that he wanted to be extubated quickly. Our team felt a sense of urgency on his behalf and was trying to balance the ethical issues of beneficence and nonmaleficence with our duty to support Joseph's mother. She needed more time to process his decision and would have to live with the memory of how Joseph's life had ended.

As a pediatric palliative care team, we increasingly see older adolescents and young adults facing critical end-of-life issues and decisions. Our biggest challenge comes when the quality of a life is debatable. Joseph's family loved him and valued his full humanity despite all of his physical limitations. Was his family's reluctance to discuss end-of-life care reflective of their belief that life in any condition is worth living? When we ask families to make decisions that involve assessing the quality of life and balancing suffering against a longer life, we are asking a question that is often difficult to hear. Making decisions based on goals of care may be more manageable, but sometimes the conflicting goals of the patient and the family have no clear and obvious priority.

There are many factors that families consider when faced with complicated and emotionally charged decisions. We have experienced a range of views that families bring to us concerning their spirituality and religious beliefs regarding medical care. Joseph's mother wanted to leave it all "up to God."

Joseph's mother wanted to leave it all "up to God."

To argue against another's belief system can alienate a patient or family member further away from the healthcare team. It is important to imagine the spiritual struggle of Joseph's mother. Why did God grant her a child only to prematurely take him away? Joseph's mother would have to go on living without him, and the story of how Joseph died would become a central element in how she would regard her life. Joseph's mother relied on her spirituality as a resource for coping. She was able to articulate a belief in an omnipotent God who had direct impact on the events in her and her son's lives. When faced with this situation, Joseph's mother's theological frame conveyed a certain resilience and adaptation for coping. Neither she nor Joseph were abandoned by God.

Had Joseph been given a tracheostomy, he would have required 24-hour nursing care at a minimum. Whereas prior to this admission Joseph already needed assistance in nearly every aspect of living, his care after a tracheostomy would likely have required admission to a nursing home, as it would have been too hard to provide the level of care he needed at home. The additional necessary care would be burdensome to Joseph as well as his caregivers. Residence in a nursing home would affect his quality of life, and those additional burdens had to be considered in the decision-making process. One of Joseph's priorities was to remain with his family, and it was likely that a tracheostomy would not allow him to reside with them.

When the team spoke directly to Joseph, it seemed that Joseph understood what was being asked of him. In the team's estimation, Joseph had the capacity to make decisions, and his desire not to receive a tracheostomy and be supported by artificial ventilation was reasonable.

> *When the team spoke directly to Joseph about the decisions he faced, it seemed he understood what was being asked of him. In the team's estimation, his desire not to undergo a tracheostomy and receive lifelong artificial ventilation was reasonable.*

But several questions remain: How do we engage a parent who is unwilling or unable to participate in discussions about end-of-life care during an acute crisis? How do we encourage families to enter these conversations before the crisis occurs? Should we have approached Joseph directly when he was an outpatient, against his mother's wishes? Because Joseph was so dependent on his mother, it would have been difficult to get him alone without his mother's consent or knowledge, and he may have been unwilling to discuss end-of-life issues without his mother. Because Joseph had reached the age of maturity and was thus an adult with capacity, the medical team had a legal and ethical responsibility to discuss his care with him. He should have been approached about how he wanted to make decisions and offered the option to create an advance directive and assign a healthcare proxy.

Joseph is not the first young man to die from DMD. He is also not the first in our PICU to die of DMD without advance care planning. His case is fairly typical of what we have experienced and expect to experience again. It is always hard to know what is best for each individual and each family, but compassion, communication, and experience can be extraordinarily helpful.

DISCLOSURES

Potential conflict of interest: Nothing to report.

BIBLIOGRAPHY

Birnkrant DJ. New challenges in the management of prolonged survivors of pediatric neuromuscular diseases: a pulmonologist's perspective. *Pediatr Pulmonol* 2006; 41: 1113–1117.

Erby LH, Rushton C, Geller G. "My son is still walking": stages of receptivity to discussions of advance care planning among parents of sons with Duchenne muscular dystrophy. *Semin Pediatr Neurol* 2006; 13: 132–140.

Feudtner C, Christakis DA, Zimmerman FJ, et al. Characteristics of deaths occurring in children's hospitals: implications of support care services. *Pediatrics* 2002; 109: 887–893.

Hilton T, Orr RD, Perkin RM, Ashwal S. End of life care in Duchenne muscular dystrophy. *Pediatr Neurol* 1993; 9: 165–177.

Lyon ME, Garvie PA, McCarter R, et al. Who will speak for me? Improving end-of-life decision-making for adolescents with HIV and their families. *Pediatrics* 2009; 123: e199–e206.

Parker D, Maddocks I, Stern LM. The role of palliative care in advanced muscular dystrophy and spinal muscular atrophy. *J Paediatr Child Health* 1999; 35: 245–250.

Pousset G, Bilsen J, De Wilde J, et al. Attitudes of adolescent cancer survivors toward end-of-life decisions for minors. *Pediatrics* 2009; 124: e1142–e1148.

Simonds AK. Respiratory support for the severely handicapped child with neuromuscular disease: ethics and practicality. *Semin Respir Crit Care Med* 2007; 28: 342–354.

Wiener L, Ballard E, Brennan T, et al. How I wish to be remembered: the use of an advance care planning document in adolescent and young adult populations. *J Palliat Med* 2008; 11: 1309–1313.

SPECIAL FEATURE

Mammographic Screening for Breast Cancer: 2010

Laurie Margolies, MD, FACR

Mount Sinai School of Medicine, New York, NY

OUTLINE

Late Breaking News

ABSTRACT

Recently, the US Preventive Services Task Force (USPSTF) guidelines concerning analog screening mammography were revised. Despite the breast cancer death rate falling by 30% since 1990 and mammography receiving the lion's share of credit for the decrease, the USPSTF no longer advises routine screening mammography for women aged 40–49. Biennial screening is advised for those aged 50–74. The USPSTF did not make a recommendation regarding screening those aged ≥75. The recommendations do not cover digital mammography. It is necessary for primary care providers to understand the data so they can assist patients making decisions about breast cancer screening ***Mt Sinai J Med 77:398–404, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: breast cancer, breast cancer screening, mammography.

The US Preventive Services Task Force (USPSTF) revised its 2002 breast cancer screening guidelines in 2009 and published them in *Annals of Internal Medicine* on November 17, 2009.[1] The task force evaluated the data and commissioned a new statistical modeling study.[2] Responding to the storm of publicity, the USPSTF issued a clarifying statement 1 month later.[3] There has been extensive public scrutiny, rebuttal, and debate, including comment by US Department of Health and Human Services Secretary Kathleen Sebelius.[4] Organizations such as the American College of Radiology and the Society of Breast Imaging, along with leading breast-imaging experts, published responses to the USPSTF guidelines, including:

- Detailed ACR Statement on Ill Advised and Dangerous USPSTF Mammography Recommendations.[5]
- ACR, SBI Statement on Annals of Internal Medicine Poll Regarding Flawed USPSTF Mammography Recommendations.[6]
- Frequently Asked Questions about Mammography and the USPSTF Recommendations: A Guide for Practitioners.[7]
- Revealing Oz Behind the Curtain: USPSTF Screening Mammography Guidelines and the Hot Air Balloon.[8]
- Why the Critics of Screening Mammography Are Wrong: They Distort Data, Rely on Weak Science, But Refuse to Defend When Challenged.[9]

Annals of Internal Medicine published an editorial defending the USPSTF with an equally provocative title: When Evidence Collides With Anecdote, Politics and Emotion: Breast Cancer screening.[10]

Women and their doctors must make decisions about breast cancer screening in the midst of debates about the benefits of mammography including the sub-debates about age-related benefits and harms and the magnitude of the benefit. Questions, misinformation, and anxiety are ubiquitous. Although the American Cancer Society recommendation is clear: "(y)early mammograms are recommended starting at age 40 and continuing for as long as a woman is in good health,"[11] the USPSTF guidelines will be challenging for the primary care physician to implement. It recommends against routine screening mammography in those younger than 50. The

Address Correspondence to:

Laurie Margolies
Department of Radiology
Mount Sinai School of Medicine
New York, NY
Email: laurie.margolies@
mountsinai.org

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20194

© 2010 Mount Sinai School of Medicine

class C recommendation states that a woman in her 40s should not routinely undergo mammography, but should take into account her values regarding benefits and harms. The grade B recommendation calls for biennial exams for women aged 50–74. A recommendation for women aged ≥75 is not made due to insufficient data; this is an "I" (insufficient data) statement. Digital mammography, as opposed to film-screen mammography, also receives an I recommendation.[1]

> *Although the American Cancer Society recommendation concerning mammographic screening is clear: "(y)early mammograms are recommended starting at age 40 and continuing for as long as a woman is in good health,"[11] the guidelines of the US Preventive Services Task Force (USPSTF) will be challenging for the primary care physician to implement.*

One might ask, "Do the guidelines matter or are they merely an intellectual exercise?" The guidelines have gained significant traction and importance, perhaps beyond what even the authors imagined. Federal healthcare legislation now mandates coverage for preventive services including mammography and eliminates copays for these services. To fall into this protected category, however, preventive-services exam must have an A or B rating from the USPSTF.[1] The only breast imaging service that falls into this category is biennial screening for women aged 50–74. Despite Secretary Sebelius stating that "they (the USPSTF) do not set federal policy and they don't determine what services are covered by the federal government,"[4] will Medicare and others stop covering yearly mammograms? Will Medicare and others stop covering digital mammography? Will insurers stop covering mammograms for normal-risk women younger than 50?

The Avon Foundation, the world's largest corporate-affiliated philanthropic organization that is focused on breast cancer and domestic violence, sought to answer this question, and recently surveyed its grant recipients from 48 states and the District of Columbia. Avon found that 25% of states, including New York, have made changes to their Breast and Cervical Cancer Early Detection programs. These programs performed fewer mammograms or eliminated screening for those younger than 50.[13] The indigent women who rely on these free breast cancer screenings will be more likely to present with interval cancers or larger cancers at diagnosis. The USPSTF will probably not include this data in subsequent analysis because it will be an observational study, not a randomized trial.

Annals of Internal Medicine conducted an online survey regarding reaction to the new guidelines and found that ≥50% of responding physicians "would definitely or probably change how they advise patients about breast cancer screening," and <25% would ignore the new guidelines. Sixty-seven percent will no longer make mammograms routine for patients aged 40–49, and 62% will now advise biennial rather than yearly mammograms for women aged 50–74. Slightly more than half will no longer obtain routine screening mammograms for women older than 75.[10]

Selection bias limits the generalizability of each of the above surveys. When all the 2010 mammography screening data are collected, one will likely see a measurable effect on mammography utilization rates. Anecdotally, many mammography practices are noting a considerable decrease in screening mammograms. It is too early to determine the lasting impact of the 2009 USPSTF guidelines on mammography screening rates and, ultimately, on breast cancer morbidity and mortality.

While the debate rages on in medical journals and the lay press, it is incumbent upon primary care physicians to understand the debate so they can assist women in making the decision about whether to have a screening mammogram. Only when a physician understands the mammography data and its analysis can he or she provide meaningful guidance to the individual patient. Daniel Kopans, MD, professor of radiology at Harvard Medical School, implores "us all to learn and understand the data and to argue in support of science and scientifically supportable data analysis."[9] Kopans, and colleagues Berg, Hendrick, and Smith, have written an excellent guide for the primary care physician: "Frequently Asked Questions about Mammography and the USPSTF Recommendations: A Guide for Practitioners."[7] I urge all to access that publication for a comprehensive understanding of the issues prior to changing a practice pattern that has worked for many years.

Mammography works. Secretary Sebelius summarized the current state of knowledge this way: "My message to women is simple. Mammograms have always been an important life-saving tool in the fight against breast cancer and they still are today."[4] Mammography usually finds breast cancer before it can be felt and has been proven to save lives. Since mammography came into widespread use, the mortality from breast cancer has fallen 30%.[9] If one looks at those women dying from breast cancer, the overwhelming majority either never have had a mammogram or have not had one in the past 2 years.[14]

> *US Department of Health and Human Services Secretary Kathleen Sebelius summarized the current state of knowledge this way: "My message to women is simple. Mammograms have always been an important life-saving tool in the fight against breast cancer and they still are today."[4] Since mammography came into widespread use, the mortality from breast cancer has fallen 30%.[9] If one looks at those women dying from breast cancer, the overwhelming majority either never have had a mammogram or have not had one in the past 2 years.[14]*

The randomized controlled studies of mammography, the only studies that the USPSTF was willing to consider, are older studies and do not account for the improved detection of cancers in young women with current technique–improved analog technique or digital technique. The studies are predominantly pre–Mammography Quality Standard Act (MQSA), and many contain 1-view mammograms, which are not used in the United States (see Figure 1). The American College of Radiology Imaging Network, in its Digital Mammographic Imaging Screening Trial (DMIST), showed that digital mammography finds many cancers that analog mammography misses in women who are younger than 50, who have dense breasts, or who are perimenopausal or premenopausal.[15] The USPSTF analysis is only a recommendation regarding film-screen mammography, which may not even exist in many institutions or communities. More than 50% of the mammography sites in the United States currently offer digital mammography. As of August 1, 2009, there were 8711 sites in the United States providing mammography services, according to the US Food and Drug Administration's monthly MQSA scorecard; about 54% were digital, and <5% of new mammography systems purchased are analog.[16] Clearly, if the goal is to save lives and detect breast cancer early, any recommendations regarding mammography for women younger than 50 should be based upon current technology. It should be very clear that the task force has only studied analog mammography.

> *The randomized controlled studies of mammography, the only studies that the USPSTF was willing to consider, are older studies and do not account for the improved detection of cancers in young women with current technique– improved analog technique or digital technique.*

It is extremely unlikely that a randomized controlled study of screening digital mammography compared with an unscreened control group could be done ethically. A trial comparing digital mammography with no screening mammography cannot and should not be performed. It is therefore incumbent upon the USPSTF to consider observational studies and retrospective studies in an analysis of current mammography practices.

As with any screening exam, mammography is not perfect. There are false positives and false negatives. False positives cause anxiety and lead to benign breast biopsies, as noted by the USPSTF. The task force, however, did not address the anxiety and anger felt by women who have not afforded themselves of screening only to find that they have advanced breast cancer. While there might be harm in screening, there is also harm in not screening.

It seems that the USPSTF, which includes experts in the evaluation of screening, did not, however, consult breast cancer or breast imaging experts, and did not subject its analysis to scrutiny by breast specialists. Indeed, "the SBI (Society of Breast Imaging) leadership learned of the USPSTF recommendations just six days prior to their publication."[17] The USPSTF routinely publishes in *Annals of Internal Medicine*: "Over the past decade, *Annals* have peer-reviewed and published more than 50 USPSTF recommendations statements and background reviews."[10] This makes the journal and its peer-review process itself open to question. *Annals of Internal Medicine* ignored a list of scientific

Study, Year	Baseline Study Year	Enrollment Age, y	Screening Protocol		Follow-up, y
			Interval, mo	View, n	
Health Insurance Plan of Greater New York, 1986	1963	40–64	12	2	18
Canadian National Breast Screening Study-1, 2002	1980	40–49	12	2	13
Gothenburg Breast Screening trial, 2003	1982	39–59	18	1–2	12
Stockholm, 2002	1981	40–64	24–28	1	11.4
Malmö, 2002	1976–1978	45–70	18–24	1–2	11–13; 15.5
Swedish Two-County trial (2 trials), 2002; 1995	1977	40–74	24–33	1	20; 15.5
Age trial, 2006	1991	39–41	12	2	10.7

Fig 1. Studies used by the US Preventive Services Task Force (Reference 2).

concerns by Dr Kopans, a recognized world expert in screening mammography, and then wrote an editorial stating that all the concerns were "anecdotal, political, or emotional."[10] *New England Journal of Medicine* and *Journal of the American Medical Association* also declined to publish scientific challenges to the USPSTF guidelines. (Daniel Kopans, MD, personal communication, April 24, 2010).

The study commissioned by the USPSTF, which was simultaneously published in *Annals of Internal Medicine*,[2] is critical to understanding the ultimate recommendations. The 2002 recommendations included "7 randomized trials of screening mammography rated as fair quality."[2] Since then, only 1 new trial was added to the database, and it too was only of fair quality, ie, not good! In this trial, ≤70% of the women actually had screening mammography (Figure 1). The Age Trial in the United Kingdom, with a mean follow-up of 10.7 years, showed a 17% reduction in mortality for women aged 39–48 who were invited to screening. If, however, one looks at those who actually were screened, the mortality reduction is 24%. This just missed being a statistically significant difference, with the 95% confidence interval at 0.51–1.01.[18] The majority of trials show a benefit to screening (see Figure 2).

For women aged 39–49, the 8 trials used in the analysis–despite being only of "fair quality" and despite using older technology–indicated "a 15% reduction in breast cancer mortality in favor of screening."[2] However, as shown by Kopans, the USPSTF used results from the flawed unblinded Canadian National Breast Screening Study to obtain this number. When looking at women younger than 50 in other trials, such as the Gothenburg Breast Cancer Screening Trial, the Malmö Mammographic Screening Program, and the Swedish trials, one sees mortality reductions of 44%, 35%, and 29%, respectively.[19] If the USPSTF used even a 29% reduction in mortality for screening, their conclusions and recommendations would have been dramatically altered. Further, the studies used measured the differences between those who were and who were not invited to be screened–not those who were screened and those who were not. Those invited may not have been screened, and those not invited may have gone for screening mammograms anyway. The screened and unscreened groups are contaminated.[18]

The USPSTF models determine that to prevent 1 cancer death, 1904 women aged 39–49 needed to be invited for screening.[2] For women aged 50–59, 1 cancer death could be prevented if 1339 were invited

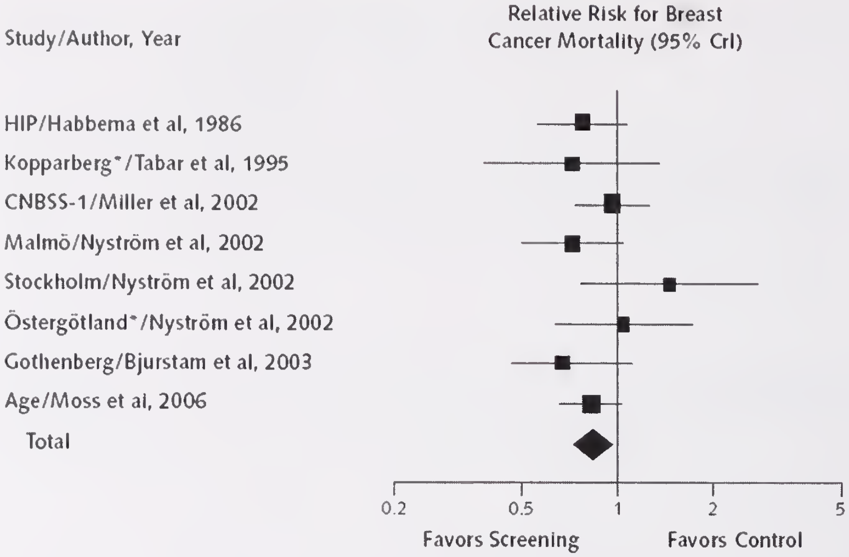


Fig 2. Relative risk for breast cancer mortality (Reference 2). **Abbreviations:** CrI, credible interval.

for screening, and for women aged 60–69, only 377 need be invited for screening to prevent 1 death.[2] So, the USPSTF's own requested study, despite all its flaws, demonstrated that screening mammography does save lives–even for those aged 39–40! Even accepting that death is an appropriate endpoint, the USPSTF shows that mammography works. A value judgment, not a scientific judgment, was made by the USPSTF.

The USPSTF's own requested study, despite all its flaws, demonstrated that screening mammography does save lives–even for those aged 39–40.

What reasons were cited for not recommending routine mammography in younger women? Potential harms cited by the study included radiation exposure, although there are no studies measuring an association between mammography radiation and breast cancer. Radiation dose is lower with modern equipment. Not only has digital mammography been shown to administer a lower radiation dose than analog mammography,[20] but newer digital mammogram units also have doses lower than those used in DMIST. Other potential harms include pain during mammography, anxiety, distress, false-positive and false-negative mammography results, additional imaging and biopsy, and overdiagnosis.[2] The analysis did not address the difference in the screening experience and results among "experts" and generalists, nor between breast centers offering rapid turnaround times and routine practices.

Overdiagnosis of breast cancer, the detection of breast cancer that would not have become clinically apparent during a woman's lifetime or that would not affect survival, has been the subject of many articles. Moss, for example, reviewed 8 screening-mammography trials and the evidence for overdiagnosis, and concluded, "Although there is a possible shift from invasive disease to DCIS (ductal carcinoma in situ), there is no evidence of overdiagnosis of all breast cancers as a result of incident screens."[21] Others used modeling to conclude "Less than 5% of cases diagnosed at prevalence screen and less than 1% of cases diagnosed at incidence screens are being overdiagnosed. Overall, we estimate overdiagnosis to be around 1% of all cases diagnosed in screened populations."[22] It is often argued that low-grade DCIS is overdiagnosed and overtreated; however, "Studies to date indicate that DCIS of all grades has the potential to progress, though high-grade lesions progress more rapidly than lower-grade lesions and are more likely to lead to metastatic disease and death."[23]

The commissioned study concludes, "Mammography screening at any age is a tradeoff of a continuum of benefits and harms. The ages at which this tradeoff becomes acceptable to individuals and society are not clearly resolved by the available evidence."[2] Evidence of mammography benefit exists–it may not, however, be in the randomized trials that use death as the endpoint.

Mammography is one of medicine's major successes and was cited as such by 225 leading general internists–not breast specialists–in 2001.[24]

I am recommending yearly mammography to all women older than 40. Based upon all the literature, I believe it to be the best practice. The death rate from breast cancer has decreased by 30% since 1990–largely due to mammographic screening. If one uses this number, the number of women needed to be screened to save 1 life is 950, which meets the USPSTF criteria.[1,2] Women should be encouraged to participate in screening with an understanding of its benefits, risks, and limitations.

Women should be encouraged to participate in screening with an understanding of its benefits, risks, and limitations.

Our energies should be focused not only on early detection, but also on studies of causation and, ultimately, prevention.

LATE BREAKING NEWS

On June 16, 2010 it was announced that the AMA resolved to recommend annual mammograms beginning at age 40. The AMA also called for the US Preventative Services Task Force to include subject matter experts on its panels and clarify its' process.[25]

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. US Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009; 151: 716–726.
2. Nelson HD, Tyne K, Naik A, et al. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2009; 151: 727–737.
3. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Material available at http://www.annals.org/content/151/10/727.full. Accessed on June 15, 2010.
4. US Department of Health and Human Services. Secretary Sebelius statement on new breast cancer recommendations. http://www.hhs.gov/news/press/2009pres/11/20091118a.html. Accessed March 20, 2010.
5. American College of Radiology. Detailed ACR statement on ill advised and dangerous USPSTF mammography recommendations. http://www.acr.org/MainMenuCategories/media_room/FeaturedCategories/PressReleases/UPSTFDetails.aspx. Accessed March 20, 2010.
6. American College of Radiology. ACR, SBI statement on Annals of Internal Medicine poll regarding flawed USPSTF mammography recommendations. http://www.acr.org/SecondaryMainMenuCategories/NewsPublications/FeaturedCategories/CurrentACRNews/archive/ACRSBIonAIMUSPSTFPoll.aspx. Accessed March 20, 2010.
7. Berg WA, Hendrick RE, Kopans DB, Smith RA. Frequently asked questions about mammography and the USPSTF recommendations: a guide for practitioners. Society of Breast Imaging, December 11, 2009. http://www.sbionline.org/associations/8199/files/Detailed_Response_to_USPSTF_Guidelines-12-11-09-Berg.pdf. Accessed March 20, 2010.
8. Javitt MC, Hendrick RE. Revealing Oz behind the curtain: USPSTF screening mammography guidelines and the hot air balloon. *AJR Am J Roentgenol* 2010; 194: 289–290.
9. Kopans D. Why the critics of screening mammography are wrong: they distort data, rely on weak science, but refuse to defend when challenged. *Diagn Imaging* 2009; 31: 1–7.
10. When evidence collides with anecdote, politics and emotion: breast cancer screening. *Ann Intern Med* 2010; 152: 531–532.
11. American Cancer Society. Guidelines for the early detection of cancer. http://www.cancer.org/docroot/ped/content/ped_2_3x_acs_cancer_detection_guidelines_36.asp. Accessed February 15, 2010.
12. Alderman L. For many, health law offers a chance for preventive care. *New York Times*. April 10, 2010; B6.
13. Avon Foundation. Survey finds access to mammograms declining for women under age 50. http://www.avoncompany.com/women/news/press20100222.html. Published February 22, 2010. Accessed April 1, 2010.
14. Cady B, Webb M, Michaelson J, et al. Death from breast cancer occurs predominantly in women not participating in mammographic screening. 2009 Breast Cancer Symposium Abstract. http://www.asco.org/ASCOv2/Meetings/Abstracts?&vmview=abst_detail_view&confID=70&abstractID=40559. Accessed April 20, 2010.
15. Pisano ED, Gatsonis C, Hendrick E, et al; Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med* 2005; 353: 1773–1783.
16. Mahdoubi K. Hologic talks digital mammography, tomosynthesis and women's health. *Dotmed News*. http://www.dotmed.com/news/story/9937. Published August 18, 2009. Accessed March 10, 2010.
17. Evans WP. Breast cancer screening guidelines: another controversy. *SBI News Winter* 2010; 1–3.
18. Moss SM, Cuckle H, Evans A, et al. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *Lancet* 2006; 368: 2053–2060.
19. Kopans DB. The recent US Preventive Services Task Force guidelines are not supported by the scientific evidence and should be rescinded. *J Am Coll Radiol* 2010; 7: 260–264.
20. Hendrick RE, Pisano ED, Averbukh A, et al. Comparison of acquisition parameters and breast dose in digital mammography and screen-film mammography in the

American College of Radiology Imaging Network digital mammographic imaging screening trial. *AJR Am J Roentgenol* 2010; 194: 362–369.

21. Moss S. Overdiagnosis and overtreatment of breast cancer: overdiagnosis in randomized controlled trials of breast cancer screening. *Breast Cancer Res* 2005; 7: 230–234.
22. Duffy SW, Agbaje O, Tabar L, et al. Overdiagnosis and overtreatment of breast cancer: estimates of overdiagnosis from two trials of mammographic screening for breast cancer. *Breast Cancer Res* 2005; 7: 258–265.
23. Jones JL. Over diagnosis and overtreatment of breast cancer: progression of ductal carcinoma in situ: the pathological perspective. *Breast Cancer Res* 2006; 8: 204.
24. Fuchs VR, Sox HC Jr. Physicians' views of the relative importance of thirty medical innovations. *Health Aff (Millwood)* 2001; 20: 30–42.
25. American College of Radiology. ACR Leads Effort to Prevent Weakening of AMA Stance on Screening Mammography http://www.acr.org/HomePageCategories/News/ACRNewsCenter/ACR-Influence-on-AMA-Stance-on-Screening-Mammography.aspx. Accessed on June 16, 2010.

WILEY-BLACKWELL

Corporate Sales

Dedicated to supporting your communication needs

Advertising | Books | Reprints
Supplements and Sponsorship

- Through our extensive scientific, clinical and professional publications, we are able to help you reach key markets by developing customized communication plans to support your business
- We offer efficient and knowledgeable service, competitive pricing, dedicated project management and the flexibility to meet the most demanding requirements across our range of products and services
- Whether you are looking for global or localized campaigns, for strategic or tactical support, our publishing teams are knowledgeable and easy to reach in your local time zone

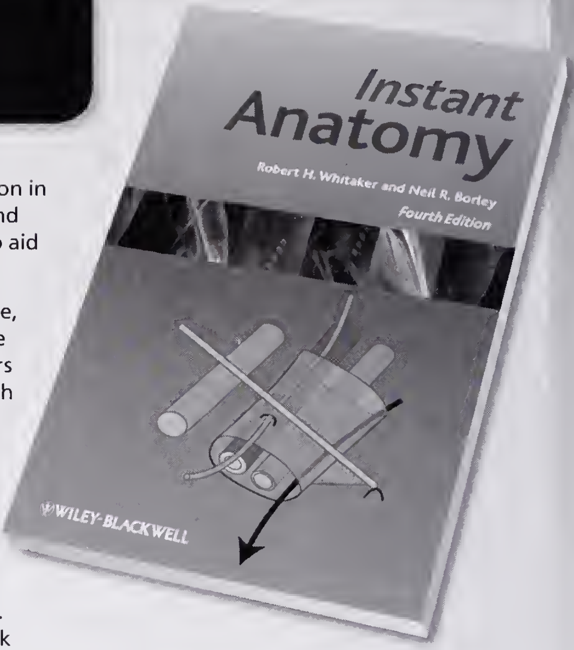
Find out more:
W: **www.wiley.com/wiley-blackwell** | E: **corporatesalesusa@wiley.com**

Europe: Chichester • Oxford • Berlin • Weinheim • Copenhagen • Madrid
North America: Hoboken • Malden | **Australia:** Melbourne | **Asia:** Tokyo • New Delhi • Shanghai • Singapore

Instant Anatomy

FOURTH EDITION

Robert H. Whitaker and Neil R. Borley
Selwyn College, University of Cambridge;
Cheltenham General Hospital, UK



Instant Anatomy uses a unique structure to provide core information in an easily accessible 'dip-in' format. It neatly catalogues anatomy and anatomical relationships in a schematic, dictionary-style manner to aid rapid understanding and retrieval of anatomical facts and data.

Now fully updated, this fourth edition of *Instant Anatomy* is precise, concise and easy to navigate, enabling quick look-up and reference points for students, junior doctors and practising clinicians. It covers all the major aspects of anatomy with a consistent, logical approach that pulls together information that would otherwise be spread between sections of a large textbook. Featuring clear illustrations and an expanded surface anatomy section, it presents complex information in a clear and understandable way.

Review of the previous editions

"*Instant Anatomy* is, without doubt, the finest book ever written... It does claim to be unique, and I think it is fair to say that, as a quick revision aid, it is unrivalled. The pictures are simple and ingenious; so much so that one wonders why nobody has written a book like this before."

Medical Student, *Murmur*
(official publication of the University of Cambridge Medical Society).

March 2010 • 9781405199612
242 pages

CONTENTS

Preface to fourth edition
Preface to first edition
Notes on the text

1: Arteries
2: Veins
3: Lymphatics
4: Autonomic nervous system
5: Cranial nerves
6: Peripheral nerves
7: Dermatomes and cutaneous nerve distribution
8: Muscles
9: Joints
10: Ossification times
11: Foramina – skull and spine
12: Position of structures according to vertebral levels
13: Pharyngeal derivatives
14: Surface anatomy and key areas

For more information or to place an order visit: **www.wiley.com**

For illustrations, podcasts, questions and answers, MCQs and PowerPoint presentations on all aspects of anatomy, visit **www.instantanatomy.net**

WILEY-BLACKWELL

The Evidence-Based Resource Trusted by Health Professionals

ESSENTIAL
EVIDENCE
PLUS

Essential Evidence Plus (EE+) is the only point-of-care resource that provides you with the Bottom Line to help your patients. This powerful new e-reference tool includes medical evidence ratings so you know you are getting the best EBM answers for teaching and patient care.

EE+ gives you everything you need to make critical patient care decisions, and includes must-have content, calculators, daily POEM updates, guidelines, and tools that clinicians can rely on for the best answer. It has the latest medical findings pre-screened for validity and relevance to clinical practice, and can be used to search all of the evidence, all of the time, in an instant. **EE+** is packed with great features you won't find with any other comparable product and is the most complete evidence-based medicine resource available!

Unique features include:

- **NEW RESOURCE!** ***Essential Evidence*** featuring 700+ current clinical topics designed to provide evidence-based answers to important questions concerning symptoms, diseases, and treatment that arise in primary care and hospital practice
- 1,900+ diagnostic test calculators
- 1,700+ history and physical test calculators
- 300+ interactive decision support tools and clinical decision rules
- Daily POEM e-mail alerts to help you stay abreast of the latest and most relevant medical literature and 3,500+ Archived POEMs
- 3,300+ Cochrane Abstracts
- EBM Guidelines brought to you by the Finnish Medical Society with 1,000+ practice guidelines, 3,000+ Evidence-Graded Summaries and 950+ high quality photographs
- Derm Expert viewer with 950+ pictures designed to help clinicians diagnose dermatological conditions
- 1,600+ Guideline Summaries
- 1,500+ Commonly used ICD-9 Codes
- Weekly Podcasts
- CME credit availability for qualified searches

Don't take our word for it.
Check out what others are saying
about Essential Evidence Plus!

"**Essential Evidence Plus** is an important tool that will provide primary care clinicians with access to the most extensive source of evidence-based medical information currently available for use at the point of care."

— Michael Wilkes, MD,
Professor of Medicine and
Vice Dean of Education,
University of California, Davis

"This database will revolutionize the way librarians work with clinicians. [It] has set the standard for all clinical database designers to surpass if they want to succeed."

— Rick Wallace, Outreach Librarian,
Eastern Tennessee State University

"...places highly valid, highly relevant information 'at the fingertips' of clinicians while they practice."

— Journal of the Medical
Library Association

"This is the most robust, able medical resource available."

— David Stockwell, MD,
Children's National Medical Center

Sign up for a free trial or subscribe today!
Group price quotations are also available upon request.

www.essentialevidenceplus.com

HEED

HEALTH ECONOMIC EVALUATIONS DATABASE

Ensures you make informed decisions about medicines and other healthcare interventions

HEED, the **Health Economics Evaluations Database,** equips health professionals with the tools they need to make informed decisions about medicines and other healthcare interventions. It does so by selecting and reviewing published economic studies which provide comparative analysis of costs and consequences.

Key Features of HEED:

- **Comprehensive coverage** – over 42,000 record (c. 21,000 full reviews and 21,000 bibliographic details only) dating from 1958-2010 with over 5,000 journals searched every week via Medline or EMBASE.
- **Up-to-date** – 250 records added each month.
- **Current** – over 300 articles published in 2010 were reviewed on HEED in the first 4 months of 2010.
- **International** – over 1,200 non-English language articles referenced.
- **Expert** – c. 1,400 non-English articles referenced.
- **Structured** – easily searchable format covering all types of economic evaluation, technology assessed, cost data, ICD codes, ATC Classifications, controlled keywords and a summary of key results.
- **Broad-range of reviews** - including cost-effectiveness analyses, cost-utility analyses, cost benefit analyses, methodology papers and many more.
- **Reliable** – the HEED selection process is a rigorous and exacting procedure that produces comprehensive coverage of economic evaluations.

WILEY-BLACKWELL

How to Order

HEED is available online on a subscription basis from Wiley-Blackwell.

For more information or to request a quote contact your Wiley Account Manager or **email heed@wiley.co.uk**.

http://heed.wiley.com

13542

VOL. 77, NO. 5
SEPTEMBER/OCTOBER 2010

MOUNT SINAI JOURNAL OF MEDICINE

A Journal of Translational and Personalized Medicine

Issue Theme
The Obesity Pandemic: Meeting the Clinical Challenges

Theme Editors
Robert Yanagisawa, MD
Derek LeRoith, MD

Medical Therapy for Obesity

Surgical Therapy for Diabesity

Metabolic Management after Bariatric Surgery

Hormone Changes after Metabolic Surgery

ffice-based Management of besity

sychological Treatments for besity

ietary Approaches to Obesity

Metabolic Aspects of Obesity: A Personalized Approach

Insulin Resistance in besity as Cause of Metabolic Syndrome

Prevention of Diabetes in Metabolic Syndrome

Weight Management in Type 2 Diabetes

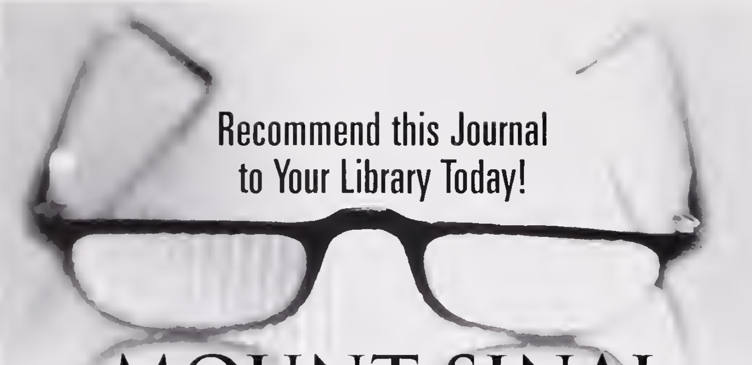
Special Features

Mechanisms of Yogic Practices in Health, Aging, and Disease

Anti-infective Lock Solutions for Catheter-related Infections

WILEY-BLACKWELL

Editor-in-Chief: Penny A. Asbell, MD, FACS, MBA



Recommend this Journal
to Your Library Today!

MOUNT SINAI
JOURNAL OF
MEDICINE

A Journal of Translational and Personalized Medicine

Editor-in-Chief: **Penny A. Asbell, MD, FACS, MBA**
Mount Sinai School of Medicine, New York, NY

Continuously published since 1934 and included in the *Index Medicus* from the very first volume, the **Mount Sinai Journal of Medicine (MSJM)** has a distinguished career as a general medical journal for a worldwide audience.

Dedicated to improving communication among all participants in medicine today, **MSJM** is an excellent source of up-to-date medical information for healthcare clinicians, educators, researchers and policy makers worldwide. **MSJM** will inspire the creative thinking needed to create new ways to understand, prevent, and treat human disease.

Each issue includes several quality, in-depth review articles related to a single theme and several others on subjects of diverse clinical interest. All theme articles are written by renowned authorities in their fields, from both Mount Sinai and elsewhere, under the supervision of Theme Editors.

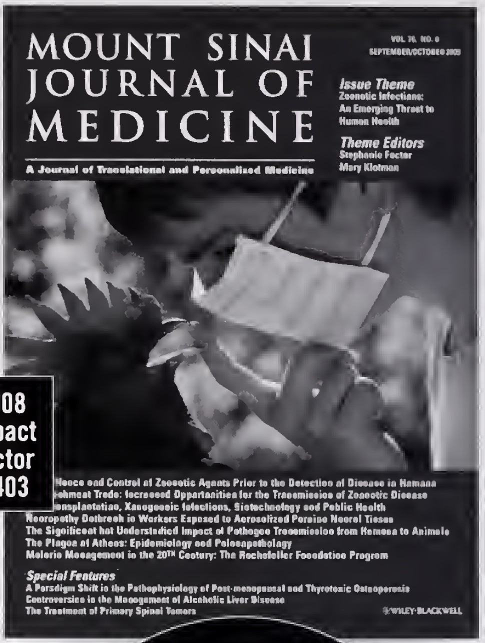
Upcoming Themes for Volume 77:

- Alzheimer's Disease: New Light on Etiology and Management
- Genes, Race, Ethnicity and the Environment
- Endovascular Surgery: Advances in Science and Technique
- Drug Discovery in Academic Settings
- The Obesity Pandemic
- Children's Health and the Environment

ACCESS THESE 2009 EDITORS CHOICE ARTICLES FREE ONLINE

- Advances and Controversies in Assisted Reproductive Technology (76:6)
- Xenotransplantation, Xenogenic Infection, Biotechnology and Public Health (76:5)
- The Student-run Health Clinic: A Novel Arena to Educate Medical Students on Systems-based Practice (76:4)
- Glycemic Targets for Patients with Type 2 Diabetes Mellitus (76:3)
- Neurobehavioral Sequelae of Traumatic Brain Injury (76:2)
- Spectrum of Fibrosing Diffuse Parenchymal Lung Disease (76:1)

Want more? Access a free sample issue online today!



MOUNT SINAI
JOURNAL OF
MEDICINE

A Journal of Translational and Personalized Medicine

Issue Theme
Zoonotic Infections:
An Emerging Threat to
Human Health

Theme Editors
Stephanie Factor
Mary Klotman

Special Features

WILEY-BLACKWELL

2008
Impact
Factor
1.403

Stay on Top
of the Research!

Sign-up for Email Table of Contents Alerts delivered direct to your desktop – FREE!

Simply go to the homepage and click on "Set E-Mail Alert" for automatic table of contents alerts, or create a customized alert based on your own search criteria.

How to Receive Complimentary Online Access for Institutions?

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Library Recommendation Form

To____________________________________

Date__________________________________

Requested by:____________________________

Dept.__________________________________

I have reviewed the **Mount Sinai Journal of Medicine** and urge you to add a complimentary 2010 online subscription to our subscription list. Please contact me should you have any questions.

WWW.MOUNTSINAIJOURNAL.ORG

WILEY-BLACKWELL

Enter Your 2010 Subscription Today!

MOUNT SINAI JOURNAL OF MEDICINE

Issue Theme
Type 2 Diabetes Mellitus: New Evidence, More Controversy

Theme Editor
Dina Green

A Journal of Translational and Personalized Medicine

Special Features

ımed at clinicians and educators, the ʼount Sinai Journal of Medicine provides gh-quality, in-depth review articles in ʼery issue, studded with clinical pearls.

ıur 2010 will include the six 2010 print sues of the **Mount Sinai Journal of edicine** (Volume 77), plus concurrent ll-text online access to these six issues ı Wiley's Internet publishing platform.

tandard Personal Subscriptions

30.00 (U.S./Canada/Mexico)
20.00 (Outside North America)

ount Sinai Alumni Subscriptions!

0.00 (Worldwide)

tended for Institutions in 2010

omplimentary nline Access!

ur institution can opt to receive mplimentary online access to the **ount Sinai Journal of Medicine** during 10. When registering for complimentary line access, your institution can also tain one complimentary print bscription.

PERSONAL SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 personal subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $105.00 (U.S./Canada/Mexico) ☐ $130.00 (Outside North America)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

Method of Payment (Canada residents please add GST. Personal rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **John Wiley & Sons, Inc.** Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA

Card #______________________________ Exp. Date______________

Signature____
(credit card orders not valid without signature)

Ship To:

Name____

Address____

City__________________________ State__________ Zip______________

MSJM0910

MSJM ALUMNI SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 Alumna/Alumnus Subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $40.00 (Worldwide)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

METHOD OF PAYMENT (Alumna/Alumnus rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **The Mount Sinai School of Medicine**. Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA ☐ Other

Card #______________________________ Exp. Date______________

Signature____
(credit card orders not valid without signature)

Ship To:

Name____

Address____

City__________________________ State__________ Zip______________

To subscribe online visit: **www.msjmonline.com**

MSJM10

COMPLIMENTARY ONLINE ACCESS FOR INSTITUTIONS

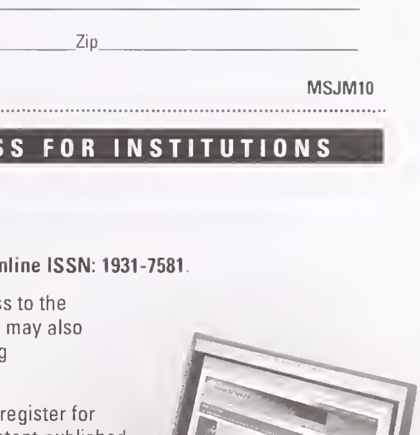
Mount Sinai Journal of Medicine

A Journal of Translational and Personalized Medicine

2010. Volume 77. 6 issues per year. Print ISSN: 0027-2507. Online ISSN: 1931-7581.

Your institution can opt to receive complimentary online access to the **Mount Sinai Journal of Medicine** during 2010. Your institution may also request one complimentary print subscription when registering for complimentary online access.

There is no obligation to subscribe in subsequent years if you register for complimentary access. Institutions retain online access to content published during the complimentary period, with or without a paid subscription in subsequent years.



Dear Librarian,
I have reviewed the scope of Mount Sinai Journal of Medicine.
Please register for complimentary access:

Name: ______________________ Department: ______________

Signature: ______________________ Date: ______________

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Don't delay, recommend to your library today!

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS MAIL PERMIT NO. 2277 HOBOKEN, NJ

POSTAGE WILL BE PAID BY ADDRESSEE

JOHN WILEY & SONS INC
ATTN: BETH MOLD, MAIL STOP 8-01
111 RIVER STREET
HOBOKEN, NJ 07030-9441

PLACE
POSTAGE
STAMP
HERE

INES SANTOS
MOUNT SINAI JOURNAL OF MEDICINE
ONE GUSTAV L. LEVY PLACE BOX 1094
NEW YORK, NY 10029-6574

Library Routing Card

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

Editor-in-Chief

Penny A. Asbell

Deputy Editor

David H. Adams

Associate Editors

Stuart Aaronson
Erwin Bottinger
Dennis S. Charney
Valentin Fuster
Wayne Goodman
Jonathan Halperin
Ravi Iyengar
Ihor R. Lemischka
Michael L. Marin
Thomas G. McGinn
Ira Nash
Maria New
Rosamond Rhodes
Hugh Sampson
Mary Sano
Jeffrey H. Silverstein
Albert Siu
Anthony Squire
Fred Suchy
Savio Woo

Editorial Board

William J. Bremner	University of Washington
Thomas M. Coffman	Duke University
Dwight Evans	University of Pennsylvania
Gabriel Haddad	University of California, San Diego
Heidi E. Hamm	Vanderbilt University
Matti Häyry	University of Manchester
Eve J. Higginbotham	Howard University
Katherine A. High	University of Pennsylvania
Tony Hope	University of Oxford
Mary Klotman	Duke University
Paul Klotman	Baylor College of Medicine
John H. Krystal	Yale University
Raju Kucherlapati	Harvard Medical School
Robert J. Lefkowitz	Duke University
Robert Lisak	Wayne State University
George Lister	University of Texas
Philip A. Marsden	University of Toronto
Charles Nemeroff	Emory University
Scott Rauch	Harvard Medical School
Bernard Roizman	University of Chicago
Alan F. Schatzberg	Stanford University
Inder Verma	The Salk Institute
Jeffrey A. Whitsett	University of Cincinnati
Allan Wolkoff	Albert Einstein College of Medicine
Noam Zohar	Bar-Ilan University

Theme Editors

Robert T. Yanagisawa, MD
Associate Professor

Derek LeRoith, MD, PhD
Professor of Molecular Medicine

Division of Endocrinology, Diabetes, and Bone Diseases
Mount Sinai School of Medicine, New York, NY

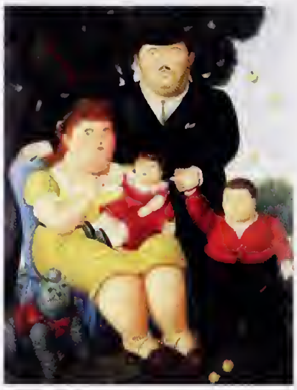
Editorial Director

Stephen Cohen
stephen.cohen@mssm.edu
212-241-6304

Publisher

John Wiley & Sons
Shawn Morton
Journal Publishing Director, Medicine
smorton@wiley.com

The Obesity Pandemic: Meeting the Clinical Challenges



"La Familia" by Fernando Botero

© Fernando Botero, courtesy Marlborough Gallery, New York, NY

Future *MSJM* Themes and Theme Editors

- Obesity pandemic (Robert Yanagisawa, Derek LeRoith)
- Translational approaches to cancer treatment (William Oh, Steven Burakoff)
- Future of primary care (Thomas McGinn)
- Patient safety (Ira Nash)
- Children's environmental health (Philip Landrigan, Amir Miodovnik)
- Biomarkers in clinical practice (Penny Asbell, Mark Babyatsky)

©2010 **Mount Sinai School of Medicine.** All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of t 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 0192 Phone: (978) 750-8400; Fax: (978) 750-4770. Such requests and other permission inquiries should be addressed to the Permissions Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboke NJ 07030; Phone: (201) 748-6011; Fax: (201) 748-6008; E-mail: permreq@wiley.com. Printed in the United States of America by The Sheridan Press.

Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine (Print ISSN: 0027-2507; Online ISSN: 1931-7581 at Wiley InterScienc www.interscience.wiley.com) is published twelve times per year for the Mount Sinai School of Medicine by Wiley Subscription Services, a Wiley Company, 111 River Street, Hoboken, NJ 07030.

Send subscription inquiries to: Subscription Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-6645.

Offprint sales and inquiries should be addressed to the Reprint Billing Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-8776. All other inquiries shou be addressed to the Customer Service Department at (201) 748-6645.

Individual subscription price: $95.00 in North America, $120.00 outside North America. Payment must be made in US dollars drawn on a US bank. Claims for undelivered copies will be accep ed only after the following issue has been delivered. For more subscription prices and information, please visit www.mountsinaijournal.org.

Change of address: Please forward six weeks prior to move: present mailing label with change of address. Please enclose a copy of the mailing label or cite your subscriber reference number order to expedite handling. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits. Send claims in care of: *Mount Sinai Journal of Medicine*, c John Wiley & Sons, Inc., Attn: Journals Admin Dept. UK, 111 River Street, Hoboken, NJ 07030.

Cancellations: Subscription cancellations will not be accepted after the first issue has been mailed. Periodicals postage paid at Hoboken, NJ and at additional mailing offices.

Postmaster: Send address changes to *Mount Sinai Journal of Medicine*, Subscription Distribution, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Printed in the United Stat of America. Single issues from current and recent volumes are available at the current single issue price from cs-journals@wiley.com. Earlier issues may be obtained from Periodicals Service Compa 11 Main Street, Germantown, NY 12526, USA. Tel: +1 518 537 4700, Fax: +1 518 537 5899, Email: psc@periodicals.com

Advertising inquiries should be addressed to: Advertising Sales Manager, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Phone: (201) 748-6921; Fax: (201) 748-62 Advertising in this journal is published with a caveat emptor understanding. The Publisher, Editors, and members of the Editorial Board do not imply warranty of products, nor quality, validity, approval of education material offered by such advertising.

Indexed by: Biological Abstracts® (Thomson ISI), Chemical Abstracts Service/SciFinder (ACS), Current Contents®/Clinical Medicine (Thomson ISI), EMBASE/Excerpta Medica (Elsevier), Ind Medicus/MEDLINE/PubMed (NLM), Science Citation Index Expanded™ (Thomson ISI), SCOPUS (Elsevier), Web of Science® (Thomson ISI).

This journal is printed on acid-free paper.

Disclaimer: All articles published, including but not limited to original research, reviews, clinical notes, clinical pearls, prismatic cases, clinician's corner, editorials, reports, and letters, represe the opinions and views of the authors and do not reflect any official policy or medical opinion of the Mount Sinai School of Medicine, the institutions with which the authors are affiliated, or t Publisher unless this is clearly specified. Articles published herein should not be relied upon as recommending or promoting a specific method, diagnosis, or treatment by physicians for a particular patient or disorder. Although the Editors and Publisher believe that drug selections and dosages and the specifications and usage of equipment and devices as set forth herein are in acco with current recommendations and practice at the time of publication, they accept no legal responsibility for any errors or omissions, and make no warranty, express or implied, with respect to ma rial contained herein.

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

CONTENTS
VOL 77, NO 5

Theme Introduction

405 The Obesity Pandemic: Meeting the Clinical Challenges
Robert Yanagisawa, Derek LeRoith

Theme Articles

407 Medical Therapy for Obesity
George A. Bray

418 Surgical Therapy for Diabesity
Kervin Arroyo, Subhash U. Kini, John E. Harvey, Daniel M. Herron

431 Metabolic Management after Bariatric Surgery
Erika Strohmayer, Michael A. Via, Robert Yanagisawa

446 Hormone Changes Affecting Energy Homeostasis after Metabolic Surgery
John Harvey, Kervin Arroyo, Judith Korner, William B. Inabnet

466 Office-based Management of Obesity
Ken Fujioka, Nitasha Bakhru

472 Psychological Treatments for Obesity
Rebecca Lo Presti, Justine Lai, Tom Hildebrandt, Katharine L. Loeb

488 Dietary Approaches to Obesity
Gal Dubnov-Raz, Elliot M. Berry

499 Personalized Approach to Metabolic Aspects of Obesity
Stephen A. Brietzke

511 Insulin Resistance in Obesity as Underlying Cause of Metabolic Syndrome
Emily J. Gallagher, Derek LeRoith, Eddy Karnieli

524 Therapeutic Options for the Prevention of Diabetes in Metabolic Syndrome
Ivana Lukacova-Zib,Geetha Gopalakrishnan

533 Weight Management in Type 2 Diabetes Mellitus
Amulya T. Siram, Robert Yanagisawa, Maria Skamagas

Special Features

549 Anti-infective Lock Solutions for Catheter-related Bloodstream Infections
Esther Y. Kim, Patricia Saunders, Nina Yousefzadeh

559 Mechanisms of Yogic Practices in Health, Aging and Disease
Viktoriya Kuntsevich, William Bushell, Neil D. Theise

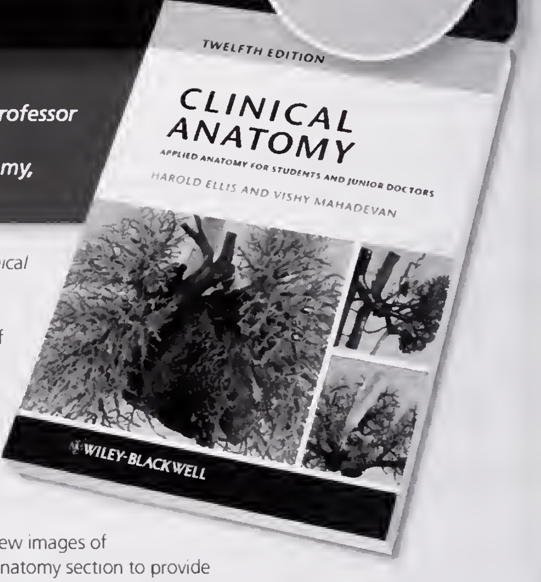
Clinical Anatomy

Applied anatomy for students and junior doctors

TWELFTH EDITION

Harold Ellis and Vishy Mahadevan

Guy's, King's and St Thomas' School of Biomedical Sciences; Emeritus Professor of Surgery, Charing Cross and Westminster Medical School, London; Barbers' Company Reader in Anatomy and Professor of Surgical Anatomy, The Royal College of Surgeons of England, London


TWELFTH EDITION
CLINICAL ANATOMY
APPLIED ANATOMY FOR STUDENTS AND JUNIOR DOCTORS
HAROLD ELLIS AND VISHY MAHADEVAN
WILEY-BLACKWELL

Written by one of the great teachers of Anatomy, this 50th anniversary edition of *Clinical Anatomy* continues to provide thousands of medical students, postgraduate trainees and junior doctors across the world with essential anatomical information within a clinical setting. It is particularly appropriate for those preparing for the Membership of the Royal College of Surgeon's Exam.

For this new edition, Professor Harold Ellis is joined, for the first time, by a co-author. Professor Vishy Mahadevan is Barber's Company Reader in Anatomy at the Royal College of Surgeons, a renowned teacher of surgical trainees and is also an examiner for the MRCS Exam. Together they provide a detailed, easy-to-follow structured text suitable for anatomy students and trainees of all levels.

Fully revised and updated, this new edition features new and re-drawn illustrations, new images of normal anatomy (radiographs, CTs and NMRs) to aid understanding, a revised neuroanatomy section to provide greater clarity and an expansion of the lower limb section.

In order to provide revision support, *Clinical Anatomy* has its own website featuring digital flashcards of all the images found in the textbook available to download for free at **www.ellisclinicalanatomy.co.uk**

March 2010 • 9781405186179
470 pages

CONTENTS

Preface to the Twelfth (Jubilee) Edition
Preface to the First Edition
Acknowledgements to the Twelfth (Jubilee) Edition
Acknowledgements to the First Edition

Part 1: The Thorax
Introduction
Surface anatomy and surface markings
The thoracic cage
The lower respiratory tract
The mediastinum
On the examination of a chest radiograph

Part 2: The Abdomen and Pelvis
Surface anatomy and surface markings
The fasciae and muscles of the abdominal wall
Peritoneal cavity
The gastrointestinal tract
The gastrointestinal adnexae: liver, gall-bladder and its ducts, pancreas and spleen
The urinary tract
The male genital organs
The bony and ligamentous pelvis
The muscles of the pelvic floor and perineum
The female genital organs
The posterior abdominal wall
Computed axial tomography

Part 3: The Upper Limb
Surface anatomy and surface markings of the upper limb
The bones and joints of the upper limb
Three important zones of the upper limb: the axilla, the cubital fossa and the carpal tunnel
The arteries of the upper limb
The brachial plexus
The course and distribution of the principal nerves of the upper limb
Compartments of the upper limb
The female breast
The anatomy of upper limb deformities
The spaces of the hand

Part 4: The Lower Limb
The surface anatomy and surface markings of the lower limb
The bones and joints of the lower limb
Three important zones of the lower limb: the femoral triangle, adductor canal and popliteal fossa
The arteries of the lower limb
The veins of the lower limb
The course and distribution of the principal nerves of the lower limb
Compartments of the lower limb

Part 5: The Head and Neck
The surface anatomy of the neck
The thyroid gland
The palate
The tongue and floor of the mouth
The pharynx
The larynx
The salivary glands
The major arteries of the head and neck
The veins of the head and neck
The lymph nodes of the neck
The cervical sympathetic trunk
The branchial system and its derivatives
The surface anatomy and surface markings of the head
The scalp
The skull
The paranasal sinuses (accessory nasal sinuses)
The mandible
The vertebral column

Part 6: The Nervous System
The brain
The spinal cord
The cranial nerves
The special senses
The autonomic nervous system

Glossary of eponyms
Index

For more information or to place an order visit:
www.ellisclinicalanatomy.co.uk

WILEY-BLACKWELL

COMING SOON!
Download the FREE digital flashcards from *Clinical Anatomy*, at **www.ellisclinicalanatomy.co.uk**

THEME INTRODUCTION

The Obesity Pandemic: Meeting the Clinical Challenges

Robert Yanagisawa, MD and Derek LeRoith, MD, PhD

Division of Endocrinology, Diabetes, and Bone Diseases, Mount Sinai School of Medicine,
New York, NY

The outbreak of obesity and the associated metabolic syndrome is continuing to spread worldwide despite our current efforts to control this epidemic. In the United States, obesity is affecting more than 1 in every 3 people, according to the Centers for Disease Control and Prevention report. Worldwide, the World Health Organization estimates that more than 1 billion adults are overweight and at least 300 million are clinically obese. In the current theme issue, "The Obesity Pandemic: Meeting the Clinical Challenges," we have compiled a multidisciplinary approach to various clinical challenges we face today and attempted to offer more personalized approaches to obesity based on clinical evidence and pathophysiology.

George Bray opens the issue with his critical review of a model for energy balance and various potential therapeutic targets for the treatment of obesity. He also offers the current options, including combination therapy, and future hopes for the optimal pharmacological therapy. Arroyo, Kini, Harvey, and Herron review the current bariatric operations, their mechanisms of action, and their ultimate effects on weight loss and control or resolution of diabetes. Consideration of nutritional, metabolic, and hormonal changes are discussed by Strohmayer, Via, and Yanagisawa to allow for optimal metabolic management following bariatric surgery. A closer look into the gut hormone changes of bariatric procedures as metabolic surgery by Harvey and Inabnet offers better understanding of incretin changes with bariatric surgery and may lead to future noninvasive treatment options targeting the root cause of metabolic syndrome.

Fujioka and Bakhru send a positive message to the practicing medical community on the importance of proper diagnosis of childhood as well as adult obesity that may help overcome barriers to the clinical care of obese individuals. As behavioral changes are the key component to treating obesity, Lo Presti, Lai, Hildebrandt, and Loeb present an important review of psychological treatments for obesity in adult and pediatric populations. Dubnov-Raz and Berry summarize the low-fat versus low-carbohydrate controversy, the potential role of a low glycemic index in weight loss, and the debate on calcium and dairy products as a strategy for losing weight. As obesity therapy should be individualized to increase its compliance and overall success, Brietzke attempts to offer a personalized approach to pathophysiology, risk assessment, and intervention.

Basic pathophysiologic mechanisms linking insulin resistance to metabolic syndrome are discussed by Gallagher, LeRoith, and Karnieli. They review complex interplay between insulin signaling, inflammation, mitochondrial function, adipokines, free fatty acids, lipid production, and glucose uptake. An overview of both lifestyle and pharmacological interventions to prevent progression of adults with metabolic syndrome to type 2 diabetes is provided by Lukacova-Zib and Gopalkrishnan. Treatment targeting the root cause of insulin resistance would not only delay the onset, but may reduce the overall risk. Siram, Yanagisawa, and Skamagas discuss the metabolic and weight effects of diabetes medications

Address Correspondence to:

Robert Yanagisawa
Division of Endocrinology,
Diabetes, and Bone Diseases,
Mount Sinai School of Medicine,
New York, NY
Email: robert.yanagisawa@
mountsinai.org

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20202

© 2010 Mount Sinai School of Medicine

in order to develop strategies for managing type 2 diabetes while helping patients maintain or lose weight in order to improve their overall health outcomes.

This issue highlights the current understanding of pathophysiology as well as ongoing clinical challenges. The authors are to be commended on their outstanding focused reviews of the various aspects of obesity. This comprehensive and multidisciplinary review will help us face the ongoing obesity epidemic with an evidence-based and more personalized approach.

DOI:10.1002/MSJ

Medical Therapy for Obesity

George A. Bray, MD, MACP

Pennington Biomedical Research Center, Louisiana State University System, Baton Rouge, LA

OUTLINE

REALITIES IN TREATING AN OBESE PATIENT
MODEL FOR ADDRESSING THE PROBLEM
DIET AND THE TREATMENT OF OBESITY
  Very-Low-Calorie Diets
  Balanced-Deficit Diets
  Low-Fat Diets
  Low-Carbohydrate Diets
  High-Protein Diets
  Comparison of Diets with Different Macronutrient Composition
MEDICATIONS AND THE TREATMENT OF OBESITY
  Mechanisms Underlying Drug Therapy of Obesity
  Drugs Approved in the United States for the Treatment of Obesity
  Drugs Used to Treat Obesity but Not Approved for This Purpose
  Drugs in Clinical Trials
CONCLUSION

ABSTRACT

Obesity results from a prolonged small positive energy imbalance, and treatment needs to reverse this imbalance. Many different diets have been tried to treat obesity, and weight loss occurs with all of them. There is currently no evidence that supports the superiority of one macronutrient composition for diets over any other. The principal effect seems to be the degree of adherence to the prescribed calorie reduction. Obesity drugs have been developed that tap brain mechanisms for controlling feeding and the gastrointestinal tract and its peptides. Orlistat blocks intestinal lipase and produces modest weight loss. Sibutramine is a serotonin-norepinephrine reuptake inhibitor that has a warning on its label from the US Food and Drug Administration because of cardiovascular risk. Its marketing has been suspended in Europe. Several drug combinations are on the horizon for treatment of obesity. ***Mt Sinai J Med 77:407–417, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: adherence, dietary treatments, drug treatments, obesity, positive energy imbalance.

Address Correspondence to:

George A. Bray
Pennington Biomedical
Research Center
Baton Rouge, LA
Email: George.Bray@pbrc.edu

The prevalence of obesity now exceeds 30% of US adults,[1] but may be leveling off at an unacceptably high level. More than two-thirds of Americans are overweight,[1] and excessive weight is a problem for children and adolescents too.[2] The fundamental problem is the result of a small, but prolonged, positive energy balance, where energy from food exceeds energy needed for everyday living.[3,4]

More than two-thirds of Americans, including children and adolescents, are overweight.

REALITIES IN TREATING AN OBESE PATIENT

One of the key messages for obese patients is that when caloric intake is reduced below that needed for daily energy expenditure, there is a predictable rate of weight loss.[4] Men generally lose weight faster than women of similar height and weight on any given diet, because men have more lean body mass and therefore higher energy expenditure. Similarly, older patients have a lower metabolic expenditure and as a rule lose weight more slowly than do younger subjects with similar adherence to weight-loss programs. Thus, adherence to any program is an essential component of success.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20207

© 2010 Mount Sinai School of Medicine

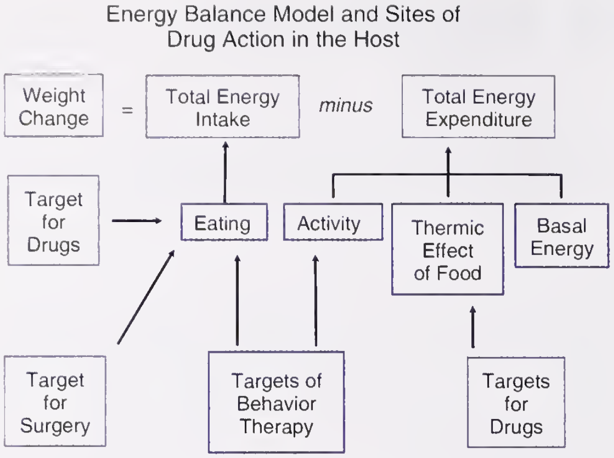


Fig 1. A model relating energy balance to treatments for obesity. The top shows that changes in energy stores are the result of changes in intake or/and energy expenditure. The other boxes show ways in which these 2 components of the system can be altered therapeutically.

MODEL FOR ADDRESSING THE PROBLEM

If obesity is the result of a prolonged small difference between energy intake and energy expenditure, then losing body fat requires reversing this imbalance. The relationship between energy intake and expenditure is shown at the top of Figure 1. Below this are various strategies that can be used to treat this imbalance.

DIET AND THE TREATMENT OF OBESITY

Reducing energy intake from food is one strategy for weight loss.[4] This can be done by reducing calories or by changing the intake of carbohydrate, fat, or protein that will in turn reduce total caloric intake. We will discuss each of these in turn.

Very-Low-Calorie Diets

Very-low-calorie diets (VLCD) or very-low-energy diets specify energy levels between 200 and 800 kcal/day. The theory behind them is that the lower the energy intake, the more rapid the weight and fat loss. Contrary to this theory are data showing that weight loss from a 400-kcal/day VLCD as compared with an 800-kcal/day diet was not different.[4] Most weight-loss diets can produce a decline in energy expenditure with a drop in triiodothyronine and leptin. Treatment with leptin partially corrects these changes.[5]

Balanced-Deficit Diets

Diets that reduce carbohydrate, protein, and fat, the so-called "balanced-deficit diets" or "prudent diets," have been widely used in treating obesity. In a meta-analysis of low-calorie diets, Avenell and colleagues[6] found that after 12 months the difference between control and treated groups was 5.31 kg (95% confidence interval [CI], −5.86 to −4.77 kg), favoring the diets. In another systematic review of 16 studies that used diet but without pharmacological intervention, and that had >100 subjects in each group and a duration of >1 year, weight loss after 2–3 years was usually <5 kg below baseline (3.5 ± 2.4 kg; range, 0.9–10.0 kg), and after 4–7 years where there were data, it was 3.6 kg ± 2.6 kg below baseline.[7]

Portion-controlled diets are one way of achieving a balanced caloric deficit. This can be done most simply by using individually packaged foods. Frozen low-calorie meals containing 250–350 kcal/package can be a convenient way to do this, except for the high salt content of many of these foods. In one 4-year study, this approach resulted in early initial weight loss, which then was maintained.[8]

> *One approach to achieve a balanced caloric deficit is a portion-controlled diet. In one 4-year study, this approach resulted in early initial weight loss, which then was maintained.*

Low-Fat Diets

Low-fat diets are a standard strategy to help patients lose weight. A meta-analysis of 5 randomized controlled trials of low-fat diets showed that these diets produced significant weight loss, but not more so than did the control diets.[9] Subsequently, the Women's Health Initiative initiated a comparison of a low-fat diet and normal dietary pattern on disease outcomes. Weight loss was a secondary outcome in this large clinical trial[10] that randomly assigned 48,835 women to either low-fat or control diets. Weight loss was 2.2 kg below baseline at year 1 and 0.6 kg after an average of 7.5 years of follow-up. At both time points the low-fat diet group was significantly lighter than the normal-fat diet group. There was a clear relationship between the decrease in percent fat and weight loss ($P < 0.001$ for trend).

Low-Carbohydrate Diets

Two strategies have been used to lower carbohydrate intake. One is to use foods with a lower glycemic index or lower glycemic load, and the other is to reduce the total amount of carbohydrate. The glycemic index is based on the rise in blood glucose in response to the test food compared with the rise after a 50-g portion of white bread; glycemic load is the product of glycemic index and amount of carbohydrate in the food.

The effect of low–glycemic index diets on weight loss has been studied in a number of randomized clinical trials in adults.[11] Thomas *et al.* identified 6 studies including 202 participants that met their inclusion criteria. Three of these studies compared low–glycemic index diets with higher–glycemic index diets, whereas the other 3 compared an ad-lib reduced–glycemic load diet with a conventional energy-restricted, reduced-fat diet, or an energy-restricted low–glycemic index diet with a normal energy-restricted diet. Interventions were relatively short, ranging from 5 weeks to 6 months at the longest. There was a small significant difference in body weight of 1.1 kg (95% CI, −2.0 to −0.2) that favored the low–glycemic index diets. The body mass decreased by 1.1 kg ($P < 0.05$) and fat mass by a similar amount compared with the change of weight in the control diet group. Both total cholesterol and low-density lipoprotein cholesterol fell more with the low–glycemic index diets.[11]

High-Protein Diets

High-protein diets have also been touted as weight-loss diets. One study compared 15%-protein and 25%-protein diets as part of a low-fat intake. Weight loss over 6 months was greater with the higher-protein diet, and this difference was maintained at 12 and 24 months. Higher-protein diets may also enhance weight maintenance.[12] Following weight loss with a VLCD for 4 weeks, 148 male and female subjects were stratified by age, body mass index (BMI), body weight, restrained eating, and resting energy expenditure, and randomized to a control condition or a supplement of 48.2 g/day of additional protein. At the end of 3 months, the group receiving the protein supplement to bring protein up to 18% had a 50% reduction in weight regain.[12]

Comparison of Diets with Different Macronutrient Composition

Several randomized clinical trials lasting 1 year[13–15] or 2 years[16,17] have compared diets head-to-head. In the first, 169 obese individuals were randomized to one of 4 popular diets: the Atkins Diet, the Ornish Diet, the Weight Watchers Diet, and the Zone Diet.[13] At the end of 12 months, each diet produced weight loss of about 5 kg, but there was no difference between diets. Adherence to the diet was the single most important criterion of success in this trial, and the Atkins and Ornish diets were apparently more difficult to adhere to. In a second 1-year trial, the Atkins, Zone, Ornish, and Lifestyle, Exercise, Attitudes, Relationships, and Nutrition (LEARN) diets were compared.[14] This trial found that the Atkins diet produced more weight loss at 6 and 12 months than the other 3 diets, which had similar results. In this study, too, a post-hoc analysis showed that adherence was the best predictor for weight loss and that the level of adherence to any diet was not very good.[18] Two reasons are proposed for the differences. First, the study by Gardner and colleagues[14] had a more homogeneous population, including only premenopausal women. Second, the Gardner study was larger and thus had more statistical power to detect differences.

The first 2-year study was conducted at a single site in Israel with a predominantly worksite population that was predominantly (83%) male.[16] There were 3 diets: a very-low-carbohydrate diet (Atkins-like), a low-fat diet, and a Mediterranean-style diet. During the first 6 months, the very-low-carbohydrate group lost the most weight, with those on the other 2 diets losing somewhat less. In the next 6 months, there was an acceleration of weight loss in the Mediterranean diet group to reach the same weight at the end of 2 years as the very-low-carbohydrate diet, with the low-fat diet group losing less weight.

The largest study, larger than the previous 3 combined, was conducted at 2 sites in the United States.[17] In total, 811 men and women were randomized to 1 of 4 diets, and 80% of them provided their weights at the end of 2 years. The diets were: (1) 20% fat and 15% protein, (2) 20% fat and 25% protein, (3) 40% fat and 15% protein, and (4) 40% fat and 25% protein. The foods used for all 4 diets were the same, only the quantities were varied. At the end of 6 months, 12 months, and 2 years, the weight loss was similar for all 4 diets. The principal determinant of success was adherence to the diet, a finding made by Dansinger *et al.*[13] and Alhassan *et al.*[18] as well.

The largest US study comparing alternative weight-loss diets involved 811 men and women over a period of 2 years. The subjects were randomized to 1 of 4 diets that involved the same foods in varying proportions. At the end of 6 months, 12 months, and 2 years, the weight loss was similar for all 4 diets. The principal determinant of success was adherence to the diet.

MEDICATIONS AND THE TREATMENT OF OBESITY

Mechanisms Underlying Drug Therapy of Obesity

Currently available medications to treat obesity work in the brain and on the gut.[4,19] A number of neurotransmitter systems, including monoamines, amino acids, and neuropeptides, are involved in modulating food intake. Serotonin 5-HT$_{2C}$ receptors modulate fat and caloric intake. Mice that cannot express the 5-HT$_{2C}$ receptor are obese and have increased food intake. Sibutramine blocks serotonin and norepinephrine reuptake. Lorcaserin, a drug in clinical trials, works directly on serotonin-2C receptors in the brain. These receptors may work through modulation of downstream melanocortin-4 receptors.[20]

Alpha$_1$ (α_1) receptors also modulate feeding. Some α_1 receptor drugs that are used to treat hypertension produce weight gain, indicating that this receptor is clinically important. In contrast, stimulation of α_2 receptors increases food intake, and a polymorphism in the α_{2a} adrenoceptor is associated with reduced metabolic rate in humans. Activation of β_2 receptors in the brain reduces food intake, and β-blocker drugs can increase body weight.

Other drugs act in the periphery. Blockade of intestinal lipase with orlistat will produce weight loss. Glucagon-like peptide 1 (GLP-1) released from intestinal L cells acts on the pancreas and brain to reduce food intake. Amylin is secreted from the pancreas and can reduce food intake.

Drugs Approved in the United States for the Treatment of Obesity

Several drugs are currently approved by the US Food and Drug Administration (FDA) for the treatment of obesity (Table 1).

Sibutramine

Sibutramine is a serotonin-norepinephrine reuptake inhibitor. In a 6-month trial in which 1047 patients were randomized to placebo or doses of sibutramine ranging from 1 to 30 mg/day, there was a clear dose-response effect.[21] Longer trials with sibutramine have been conducted in uncomplicated obese patients, in patients with hypertension, in diabetics, and in children.[4,19]

Efficacy In a meta-analysis of long-term studies of sibutramine, the placebo group had a weighted mean weight loss of 2.18 ± 5.23 kg. The drug-treated group had a weighted mean weight loss of 6.35 ± 6.47 kg, for a net, or placebo-subtracted, weight loss of 4.16 kg (95% CI, −4.73 to −3.59).[22] Sibutramine can be used as intermittent therapy.[23] During this 12-month trial, the continuous-therapy and intermittent-therapy groups lost the same amount of weight. Sibutramine has been used in children.[24] In a large 12-month multicenter trial with 498 adolescents aged 12–16 years, the mean absolute change in BMI was 2.9 kg/m^2 (8.2%) in the sibutramine group compared with 0.3 kg/m^2 (0.8%) in the placebo group ($P < 0.001$).[25] Sibutramine has also been studied as part of a behavioral weight-loss program. With sibutramine alone and minimal behavioral intervention, the weight loss over 12 months was approximately 5.0 ± 7.4 kg. Behavior modification alone produced a weight loss of 6.7 ± 7.9 kg. Adding a brief behavioral therapy session to a group that also received sibutramine produced a slightly larger weight loss of 7.5 ± 8.0 kg. When the intensive

In a meta-analysis of long-term studies of sibutramine, the placebo group had a weighted mean weight loss of 2.18 ± 5.23 kg. The drug-treated group had a weighted mean weight loss of 6.35 ± 6.47 kg, for a net, or placebo-subtracted, weight loss of 4.16 kg (95% CI, −4.73 to −3.59).

Table 1. *Drugs Approved by the US Food and Drug Administration That Produce Weight Loss*[19].

Generic Name	Trade Name(s)	Usual Dose	Comments
For long-term treatment of obesity (12 months)			
Orlistat	Xenical	120 mg, 3 times a day	May have gastrointestinal side effects
Sibutramine*	Meridia (United States), Reductil (elsewhere)	5–15 mg, once daily	Norepinephrine-serotonin reuptake inhibitor; may raise blood pressure
For short-term treatment of obesity (12 weeks)			
Diethylpropion*			Sympathomimetic drugs; approved for only a short time
Tablets	Tenuate	25 mg, 3 times a day	
Extended release	Tenuate	75 mg, in AM	
Phentermine HCl*			
Capsules	Phentridol, Tetramine, Adipex-P	15–37.5 mg, in AM	
Tablets	Tetramine, Adipex-P	15–37.5 mg, in AM	
Extended release	Ionamin	15 or 30 mg/day, in AM	
Benzphetamine†	Didrex	25–150 mg/day, in single or divided doses	
Phendimetrazine†			
Extended-release capsules	Adipost, Bontril, Melfiat, Prelu-2, X-trozine	105 mg, once daily	
Tablets	Bontril, Obezine	35 mg, 2 or 3 times a day	

*US Drug Enforcement Agency Schedule IV.
†US Drug Enforcement Agency Schedule III.

lifestyle intervention was combined with sibutramine, the weight loss increased to 12.1 ± 9.8 kg.[26]

Safety Sibutramine increases blood pressure (BP) levels in normotensive patients or blunts the decrease that might have occurred with weight loss. Systolic and diastolic BP increased an average of +0.8 mm Hg and +0.6 mm Hg, respectively, and pulse increased approximately 4–5 beats per minute.[19,26] Caution should be used when combining sibutramine with other drugs that may increase BP. Sibutramine is contraindicated in patients with a history of coronary artery disease, congestive heart failure, cardiac arrhythmias, or stroke. Sibutramine should not be used with selective serotonin reuptake inhibitors or monoamine oxidase inhibitors, and there should be a 2-week interval between terminating monoamine oxidase inhibitors and beginning sibutramine. Because sibutramine is metabolized by the cytochrome P-450 enzyme system (isozyme CYP3A4), it may interfere with the metabolism of erythromycin and ketoconazole, and it has a small effect (7% increase in area under the curve) on metabolism of simvastatin, but not other statins. An outcomes trial begun when sibutramine was approved, the Sibutramine Cardiovascular Outcomes Trial (SCOUT), has shown that there are increased deaths among those treated with sibutramine. For this reason, the marketing of this drug has been suspended in Europe.

> *An outcomes trial begun when sibutramine was approved, the Sibutramine Cardiovascular Outcomes Trial, has shown that there are increased deaths among those treated with sibutramine. For this reason, the marketing of this drug has been suspended in Europe.*

Sympathomimetic Drugs

The sympathomimetic drugs benzphetamine, diethylpropion, phendimetrazine, and phentermine are grouped together because they act like norepinephrine and were tested before 1975.

Efficacy One of the longest of the clinical trials of drugs in this group lasted 36 weeks and compared placebo treatment with continuous phentermine or intermittent phentermine.[19] Both continuous and intermittent phentermine therapy produced more weight loss than did placebo. In one survey of bariatric physicians, use of these sympathomimetic amines was found to be more frequent than sibutramine or orlistat, and they were often used for longer than approved by the FDA.[27]

Safety Sympathomimetic drugs produce insomnia, dry mouth, asthenia, and constipation. They are scheduled by the US Drug Enforcement Agency, suggesting the government's view that they may be abused. Sympathomimetic drugs can also increase BP. If a physician decides to use any of these drugs more than 3 months in succession, it would be appropriate to obtain written informed consent from the patient.

Orlistat

Orlistat is a potent and selective inhibitor of pancreatic lipase that reduces intestinal digestion of fat. It is available as a prescription drug, and at half the dose as an over-the-counter preparation.

Efficacy A number of long-term clinical trials with orlistat have been published using uncomplicated obese patients and obese patients with diabetes. A 4-year, double-blind, randomized, placebo-controlled trial with orlistat in 3304 overweight patients, 21% of whom had impaired glucose tolerance,[28] achieved a weight loss during the first year of >11% below baseline in the orlistat-treated group compared with 6% below baseline in the placebo-treated group. Over the remaining 3 years of the trial, there was a small regain in weight, such that by the end of 4 years, the orlistat-treated patients were 6.9% below baseline, compared with 4.1% for those receiving placebo. There was a reduction of 37% in the conversion of patients from impaired glucose tolerance to diabetes. Orlistat has also been studied in adolescents. In 539 adolescents, orlistat 120 mg 3 times/day decreased BMI by 0.55 kg/m^2 in the drug-treated group compared with an increase of +0.31 kg/m^2 in the placebo group.[29] In a meta-analysis of trials with orlistat, the weighted mean weight loss in the placebo group was 2.40 ± 6.99 kg and the weight loss in those treated with orlistat was 5.70 ± 7.28 kg, for a net effect of 2.87 (95% CI, −3.21 to −2.53).[22]

A 4-year, double-blind, randomized, placebo-controlled trial with orlistat in 3304 overweight patients, 21% of whom had impaired glucose tolerance, achieved a weight loss during the first year of more than 11% below baseline in the orlistat-treated group compared with 6% below baseline in the placebo-treated group.

Safety Orlistat is not absorbed to any significant degree, and its side effects are thus related to the blockade of triglyceride digestion in the intestine.[19] Fecal fat loss and related gastrointestinal symptoms are common initially, but they subside as patients learn to use the drug. Orlistat can cause small but significant decreases in fat-soluble vitamins. Levels usually remain within the normal range, but a few patients may need vitamin supplementation. Because it is impossible to tell which patients need vitamins, it is wise to provide a multivitamin routinely, with instructions to take it before bedtime. Orlistat does not seem to affect the absorption of other drugs, except acyclovir. Rare cases of severe liver injury have been reported with the use of orlistat; only 1 occurred in the United States and 13 occurred elsewhere over 10 years, during which period an estimated 40 million people took orlistat.[30] A causal relationship has not been established, but patients who take orlistat should contact their healthcare provider if itching, jaundice, pale-colored stools, or anorexia develop.

Combining Orlistat and Sibutramine

When orlistat was combined with sibutramine in a 4-month trial initiated at the end of a 12-month treatment period with sibutramine, there was no further weight loss.[31] Thus, we have no data that combining orlistat and sibutramine is beneficial.

Drugs Used to Treat Obesity but Not Approved for This Purpose

Several drugs approved by the FDA for purposes other than weight loss have been used for the treatment of obesity, including metformin, fluoxetine, bupropion, topiramate, and zonisamide.[4] Physicians who decide to try these agents are using them off-label and would be well advised to do so only with signed informed consent from the patient.[32]

Antidiabetic Drugs that Lower Body Weight

Exenatide Exenatide (Exendin-4) is a 39 amino acid peptide that is produced in the salivary gland of the Gila monster, a species of lizard native to the southwestern United States and Mexico. It has 53% homology with GLP-1, but it has a much longer half-life. Exenatide has been approved by the FDA for treatment of type 2 diabetics who are inadequately controlled while being treated with either metformin or sulfonylureas. In human beings, exenatide reduces fasting and postprandial glucose levels, slows gastric emptying, and decreases food intake by 19%.[33] A systematic review of incretin therapy in type 2 diabetes[34] showed a weight loss of 2.37 kg for all GLP-1 analogs versus control, 1.44 for exenatide versus placebo injection, and 4.76 for exenatide versus insulin (which often leads to weight gain). A 24-week, multicenter, randomized, placebo-controlled clinical trial of exenatide enrolled diabetics poorly controlled with either metformin or sulfonylurea with a HbA_{1c} between 6.6% and 10.0% and a BMI between 25.0 and 39.9 kg/m^2. The decrease in caloric intake by exenatide (378 ± 58 kcal/day) was not significantly different from placebo (295 ± 58 kcal/day). Weight loss was significantly greater and HbA_{1c} and BP were reduced more in the exenatide-treated patients.[35] The side effects of exenatide in humans are headache, nausea, and vomiting; they are lessened by gradual dose escalation.[36] The interesting feature of this weight loss is that it occurred without prescribing lifestyle modification, diet, or exercise. A 26-week randomized control trial of exenatide produced a 2.3-kg weight loss compared with a gain of 1.8 kg in the group receiving the glargine form of insulin.[37]

> *In a 24-week, multicenter, randomized, placebo-controlled clinical trial comparing exenatide with placebo in the treatment of poorly controlled diabetic patients, exenatide produced significantly greater decreases in weight, HbA_{1c}, and blood pressure. The interesting feature of the weight loss is that it occurred with no prescription for lifestyle modification, diet, or exercise.*

Liraglutide Liraglutide is another GLP-1 agonist that has a 97% homology to GLP-1. This molecular change extends the circulating half-life from 1–2 minutes to 13 hours. Liraglutide reduces body weight. In a 20-week, multicenter European clinical trial, Astrup *et al.*[38] reported that daily injections of liraglutide at 1.2, 1.8, 2.4, or 3.0 mg produced weight losses of 4.8, 5.5, 6.3, and 7.2 kg, respectively, compared with 2.8 kg in the placebo-treated group and 4.1 kg in the orlistat-treated comparator group. In the group treated with 3.0 mg/day, 76% achieved a >5% weight loss, compared with 30% in the placebo group. BP was significantly reduced, but there was no change in lipids. The prevalence of prediabetes was reduced by liraglutide. In a head-to-head comparison, liraglutide and exenatide produced similar amounts of weight loss, 3.24 kg with liraglutide versus 2.87 kg with exenatide. In poorly controlled type 2 diabetics on maximally tolerated doses of metformin and/or sulfonylurea, liraglutide reduced mean HbA_{1c} significantly more than did exenatide (−1.12% versus −0.79%).[39] Liraglutide has been approved by the European Medicines Agency and the FDA for the treatment of diabetes.

Combinations of Drugs that Produce Weight Loss

The first important clinical trial for weight loss combining drugs that acted by separate mechanisms used phentermine and fenfluramine.[40] This trial showed a highly significant weight loss of nearly 15% below baseline with fewer side effects by using combination therapy. This combination became very popular, but due to reports of aortic valvular regurgitation associated with its use,[41] fenfluramine was withdrawn from the market worldwide on September 15, 1997. Several other combinations of existing drugs are now under development and may be the next therapeutic option for treatment of obesity.

Topiramate and Phentermine Topiramate is an anticonvulsant drug that was shown to reduce food intake, but it was not developed clinically because of the side effects at the doses selected for trial. Phentermine is a long-established sympathomimetic drug approved for the short-term treatment of obesity in 1968. It is a "scheduled" drug, meaning that the US Drug Enforcement Agency has concluded that this drug carries risk for habituation and/or addiction. Topiramate and phentermine have been combined in one of 3 combinations: 3.75, 7.5, or 15 mg of phentermine combined, respectively, with 23, 46, or 92 mg of topiramate.

Two clinical trials have been conducted to evaluate long-term efficacy and safety. One is in obese individuals with a BMI ≥35 kg/m^2 and a normal cardiometabolic profile (BP <140/90 mm HG, fasting plasma glucose <110 mg/dL, triglycerides <200 mg/dL; Healthy Obese Trial). The other is in people with obesity combined with comorbidities (BMI 27–45 kg/m^2 and 2 or more of the components for the metabolic syndrome, including treatment for hypertension, diabetes, or dyslipidemia; Obesity with Comorbidities). The LEARN manual provided the behavioral program for both trials. In the Healthy Obese trial, the low and high doses of topiramate/phentermine were used. In the Obesity with Comorbidities trial, topiramate and phentermine were used at the mid and high dose. The medication was titrated over 4 weeks due to the side effects from topiramate. All together, more than 2500 patients were included in these clinical trials. By design, the Obesity with Comorbidities trial had an older population with many more comorbidities but a lower BMI than the other trial. Weight loss among those who completed the trial was similar for the high dose in both trials. The percentage of those losing >5%, >10%, or >15% was also similar between the 2 trials. At the end of 1 year, there was an improvement in all risk factors in the high-dose group, and all but diastolic BP and low-density lipoprotein cholesterol in the middle-dose group. Side effects in both trials included tingling and paresthesias. The prevalence of mental and behavioral side effects noted with topiramate alone was not as prominent in these trials. Using the intent-to-treat basis for analyzing data, the full dose of the combination produced an 11% weight loss at 1 year compared with 1.6% in the placebo-treated group.

Two clinical trials compared a combination of topiramate and phentermine with placebo in the treatment of more than 2500 obese patients with and without comorbidities. Using the intent-to-treat basis for analyzing data, the full dose of the combination produced an 11% weight loss at 1 year compared with 1.6% in the placebo-treated group.

Bupropion and Naltrexone Bupropion reduces food intake by acting on adrenergic and dopaminergic receptors in the hypothalamus. Naltrexone is an opioid receptor antagonist with minimal weight-loss effect on its own. The rationale for combining bupropion with naltrexone is the idea that bupropion, acting on adrenergic receptors, is stimulating the production of a prohormone, proopiomelanocortin, which contains several other peptides within it, including α-melanocyte stimulating hormone (α-MSH), adrenocorticotropin, and β-endorphin. The α-MSH reduces food intake, but β-endorphin stimulates food intake. Naltrexone blocks β-endorphin and increases the release of α-MSH, which can continue to inhibit feeding.[42] In a 24-week dose-ranging study, 419 subjects were randomized but only 244 (64%) completed the trial. Among the completers, weight loss was 1.2 kg in the placebo group, 3.1 kg in the bupropion-treated group, but only 1.6 kg with the highest (48 mg/day) dose of naltrexone. When combined, weight loss was greater and similar at all 3 doses of naltrexone, combined with bupropion (7.1 kg at 16 mg/day naltrexone + bupropion, 6.6 kg at 32 mg/day naltrexone + bupropion, and 6.9 kg at 48 mg/day naltrexone + bupropion). Nausea was the predominant side effect.[43]

Four main trials have been done to evaluate this combination. One trial used both 16 and 32 mg of naltrexone with 360 mg/day of bupropion, and the second used a single dose of naltrexone/bupropion (32 mg/360 mg per day), with rerandomization to 32 mg/day or 48 mg/day of naltrexone + bupropion for nonresponders at week 28. A clinical trial including behavior modification included 793 participants who were randomized in a 1:3 ratio to the combination of naltrexone 32 mg with bupropion 360 mg, provided in 2 doses per day with 2 pills at each containing 8 mg of naltrexone and 90 mg of bupropion. The dose was escalated over 4 weeks. At 56 weeks, weight loss

A randomized clinical trial compared a combination of bupropion, naltrexone, and behavior modification with placebo in the treatment of 793 obese patients. At 56 weeks, weight loss was 5.1% in the placebo group compared with 9.3% in the naltrexone 32 mg/bupropion 360 mg group.

was 5.1% in the placebo group compared with 9.3% in the naltrexone 32 mg/bupropion 360 mg group.[44] When only those who completed the trial were examined, the weight loss was 7.3% in the control group and 11.5% in the naltrexone/bupropion group. Nausea, constipation, and headache were among the more prominent side effects. There was no evidence of increased suicidal thoughts.

Pramlintide and Leptin Pramlintide is a modified form of amylin, a peptide secreted from the beta cells of the pancreas along with insulin. Pramlintide has been approved by the FDA for treatment of diabetes and produced weight loss in clinical trials. Leptin is a secretory product, primarily from adipocytes, that can act as a negative feedback signal to the brain and inhibit food intake. In clinical trials, leptin produced disappointing weight loss.[45] The combination of leptin with pramlintide, however, produced additive weight loss in a 6-month clinical trial that may offer promise for the future.[46]

> *Clinical trials have shown weight-loss results with leptin alone to be disappointing, but a 6-month trial combining leptin with pramlintide produced an additive weight loss that may offer promise for the future.*

Drugs in Clinical Trials

Lorcaserin

The neurotransmitter serotonin is involved in the regulation of food intake and food preference.[4,19] Mice lacking the serotonin 5-HT$_{2C}$ receptor have increased food intake because they take longer to become satiated. These mice are also resistant to fenfluramine, a serotonin agonist that causes weight loss. Lorcaserin is a potent, selective serotonin 5-HT$_{2C}$ agonist with ~15-fold and 100-fold selectivity versus 5-HT$_{2A}$ and 5-HT$_{2B}$ receptors, respectively. In a 12 week dose-ranging study, a total of 459 male and female subjects with a BMI between 29 and 46 kg/m^2, with an average weight of 100 kg, were enrolled in a randomized, double-blind, controlled trial comparing placebo against 10 and 15 mg given once daily and 10 mg given twice daily (20 mg/day). The placebo group lost 0.32 kg compared with 1.8 kg in the 10 mg/day dose given twice daily, 2.6 kg in the 15 mg/day dose, and 3.6 kg in the 10 mg/day dose given twice daily (20 mg total). The proportions of completers achieving ≥5% of initial body weight were 12.8%, 19.5%, 31.2%, and 2.3% in the 10-mg once daily group, 15 mg once daily, 10 mg twice daily, and placebo groups, respectively. The most frequent adverse events were transient headache, nausea, and dizziness. Echocardiograms showed no apparent drug-related effects on heart valves or pulmonary artery pressure.[47]

In a 2-year, double-blind, multicenter study, 1595 obese men and women were randomized to lorcaserin 10 mg twice daily and 1587 to placebo. Participants received a 600 kcal/day deficit diet and a healthy lifestyle program with occasional meetings with a dietitian and a recommendation to walk 30 minutes/day. Of the 883 subjects in the lorcaserin group, 716 entered the second year. Body weight reached a minimum of 9.7 kg below baseline by 48 weeks, but had risen to 6.0 kg below baseline at 102 weeks in those who remained on lorcaserin throughout. In those switched to placebo, the weight increase was greater, rising to 4.2 kg below baseline at 48 weeks, compared with 3.85 kg below baseline at 48 weeks and 2.7 kg below baseline at 102 weeks in the placebo group. BP, heart rate, triglycerides, glucose, homeostasis model assessment of insulin resistance, and C-reactive protein improved, but there was no change in high-density lipoprotein cholesterol.[48]

Tesofensine

Tesofensine is a multiamine reuptake inhibitor. In one 6-month clinical trial, this drug produced dose-related weight loss, with the highest dose producing >10% weight loss.[49]

CONCLUSION

Both diet and medications are useful in treatment of the obese patient. Weight loss of about 10% below baseline is seen with both, and there is no evidence that the composition of the diet, by itself, has any influence on weight loss. There are presently only 2 drugs approved for long-term treatment of overweight patients, and their effectiveness is limited to palliation of the chronic disease of obesity. Combinations of medications and antidiabetic drugs that produce weight loss are being evaluated.

DISCLOSURES

Dr. Bray has served as a consultant to Herbalife and VIVUS.

REFERENCES

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999–2008. *JAMA* 2010; 303: 235–241.
2. Ogden CL, Carroll MD, Curtin LR, et al. Prevalence of high body mass index in US children and adolescents, 2007–2008. *JAMA* 2010; 303: 242–249.
3. Bray GA. *The Low-Fructose Approach to Weight Control*. Pittsburgh, PA: Dorrance Publishing; 2009.
4. Bray GA. *The Metabolic Syndrome and Obesity*. Totowa, NJ: Humana Press; 2007.
5. Rosenbaum M, Goldsmith R, Bloomfield D, et al. Low-dose leptin reverses skeletal muscle, autonomic, and neuroendocrine adaptations to maintenance of reduced weight. *J Clin Invest* 2005; 115: 3579–3586.
6. Avenell A, Broom J, Brown TJ, et al. Systematic review of the long-term effects and economic consequences of treatments for obesity and implications for health improvement. *Health Technol Assess* 2004; 8: iii–iv, 1–182.
7. Douketis JD, Macie C, Thabane L, Williamson DF. Systematic review of long-term weight loss studies in obese adults: clinical significance and applicability to clinical practice. *Int J Obes (Lond)* 2005; 29: 1153–1167.
8. Flechtner-Mors M, Ditschuneit HH, Johnson TD, et al. Metabolic and weight loss effects of long-term dietary intervention in obese patients: four-year results. *Obes Res* 2000; 8: 399–402.
9. Pirozzo S, Summerbell C, Cameron C, Glasziou P. Should we recommend low-fat diets for obesity [published correction appears in *Obes Rev* 2003;4:185]. *Obes Rev* 2003; 4: 83–90.
10. Howard BV, Manson JE, Stefanick ML, et al. Low-fat dietary pattern and weight change over 7 years: the Women's Health Initiative Dietary Modification Trial. *JAMA* 2006; 295: 39–49.
11. Thomas DE, Elliott EJ, Baur L. Low glycaemic index or low glycaemic load diets for overweight and obesity. *Cochrane Database Syst Rev* 2007; 3: CD005105.
12. Paddon-Jones D, Westman E, Mattes RD, et al. Protein, weight management, and satiety. *Am J Clin Nutr* 2008; 87: 1558S–1561S.
13. Dansinger ML, Gleason JA, Griffith JL, et al. Comparison of the Atkins, Ornish, Weight Watchers, and Zone diets for weight loss and heart disease risk reduction: a randomized trial. *JAMA* 2005; 293: 43–53.
14. Gardner CD, Kiazand A, Alhassan S, et al. Comparison of the Atkins, Zone, Ornish, and LEARN diets for change in weight and related risk factors among overweight premenopausal women: the A TO Z Weight Loss Study: a randomized trial [published correction appears in *JAMA* 2007;298:178]. *JAMA* 2007; 297: 969–977.
15. Davis NJ, Tomuta N, Schechter C, et al. Comparative study of the effects of a 1-year dietary intervention of a low-carbohydrate diet versus a low-fat diet on weight and glycemic control in type 2 diabetes. *Diabetes Care* 2009; 32: 1147–1152.
16. Shai I, Schwarzfuchs D, Henkin Y, et al; Dietary Intervention Randomized Controlled Trial (DIRECT) Group. Weight loss with a low-carbohydrate, Mediterranean, or low-fat diet [published correction appears in *N Engl J Med* 2009;361:2681]. *N Engl J Med* 2008; 359: 229–241.
17. Sacks FM, Bray GA, Carey VJ, et al. Comparison of weight-loss diets with different compositions of fat, carbohydrate and protein. *N Engl J Med* 2009; 360: 859–873.
18. Alhassan S, Kim S, Bersamin A, et al. Dietary adherence and weight loss success among overweight women: results from the A to Z weight loss study. *Int J Obes* 2008; 32: 985–991.
19. Bray GA, Greenway FL. Pharmacological treatment of the overweight patient. *Pharmacol Rev* 2007; 59: 151–184.
20. Lam DD, Pryzdzial MJ, Ridley SH, et al. Serotonin 5-HT2C receptor agonist promotes hypophagia via downstream activation of melanocortin 4 receptors. *Endocrinology* 2008; 40: 1323–1328.
21. Bray GA, Blackburn GL, Ferguson JM, et al. Sibutramine produces dose-related weight loss. *Obes Res* 1999; 7: 189–198.
22. Rucker D, Padwal R, Li SK, et al. Long term pharmacotherapy for obesity and overweight: updated meta-analysis [published correction appears in *BMJ* 2007;335. DOI: 10.1136/bmj.39406.519132.AD]. *BMJ* 2007; 335: 1194–1199.
23. Wirth A, Krause J. Long-term weight loss with sibutramine: a randomized controlled trial. *JAMA* 2001; 286: 1331–1339.
24. Berkowitz RI, Fujioka K, Daniels SR, et al; Sibutramine Adolescent Study Group. Effects of sibutramine treatment in obese adolescents: a randomized trial. *Ann Intern Med* 2006; 145: 81–90.
25. Wadden TA, Berkowitz RI, Womble LG, et al. Randomized trial of lifestyle modification and pharmacotherapy for obesity. *N Engl J Med* 2005; 353: 2111–2120.
26. Kim SH, Lee YM, Jee SH, Nam CM. Effect of sibutramine on weight loss and blood pressure: a meta-analysis of controlled trials. *Obes Res* 2003; 11: 1116–1123.
27. Hendricks EJ, Rothman RB, Greenway FL. How physician obesity specialists use drugs to treat obesity. *Obesity (Silver Spring)* 2009; 17: 1730–1735.
28. Torgerson JS, Hauptman J, Boldrin MN, Sjöström L. XENical in the prevention of diabetes in obese subjects (XENDOS) study: a randomized study of orlistat as an adjunct to lifestyle changes for the prevention of type 2 diabetes in obese patients [published correction appears in *Diabetes Care* 2004;27:856]. *Diabetes Care* 2004; 27: 155–161.
29. Chanoine JP, Hampl S, Jensen C, et al. Effect of orlistat on weight and body composition in obese adolescents: a randomized controlled trial [published correction appears in *JAMA* 2005;294:1491]. *JAMA* 2005; 293: 2873–2883.
30. FDA Drug Safety Communication: Completed safety review of Xenical/Alli (orlistat) and severe liver injury. http://www.fda.gov/Drugs/DrugSafety/Postmarket DrugSafetyInformationforPatientsandProviders/ucm213038.htm. Published May 26, 2010. Accessed June 2, 2010.
31. Wadden TA, Berkowitz RI, Womble LG, et al. Effects of sibutramine plus orlistat in obese women following 1 year of treatment by sibutramine alone: a placebo-controlled trial. *Obes Res* 2000; 8: 431–437.
32. Snow V, Barry P, Fitterman N, et al. Pharmacologic and surgical management of obesity in primary care: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2005; 142: 525–531.

33. Edwards CM, Stanley SA, Davis R, et al. Exendin-4 reduces fasting and postprandial glucose and decreases energy intake in healthy volunteers. *Am J Physiol Endocrinol Metab* 2001; 281: E155–E161.
34. Amori RE, Lau J, Pittas AG. Efficacy and safety of incretin therapy in type 2 diabetes: systematic review and meta-analysis. *JAMA* 2007; 298: 194–206.
35. Heine RJ, Van Gaal LF, Johns D, et al. Exenatide versus insulin glargine in patients with suboptimally controlled type 2 diabetes: a randomized trial. *Ann Intern Med* 2005; 143: 559–569.
36. Apovian CM, Bergenstal RM, Cuddihy RM, Qu Y, Lenox S, Lewis MS, Glass LC. Effects of exenatide combined with lifestyle modification in patients with type 2 diabetes. *Am J Med* 2010 May; 123: 468.e9-17.
37. Fineman MS, Shen LZ, Taylor K, Kim DD, Baron AD. Effectiveness of progressive dose-escalation of exenatide (exendin-4) in reducing dose-limiting side effects in subjects with type 2 diabetes. *Diabetes Metab Res Rev* 2004; 20: 411–417.
38. Astrup A, Rössner S, Van Gaal L, et al.; NN8022-1807 Study Group. Effects of liraglutide in the treatment of obesity: a randomised, double-blind, placebo-controlled study [published correction appears in *Lancet* 2010;375:984]. *Lancet* 2009; 374: 1606–1616.
39. Buse JB, Rosenstock J, Sesti G, et al.; LEAD-6 Study Group. Liraglutide once a day versus exenatide twice a day for type 2 diabetes: a 26-week randomised, parallel-group, multinational, open-label trial (LEAD-6). *Lancet* 2009; 374: 39–47.
40. Weintraub M. Long-term weight control: the National Heart, Lung, and Blood Institute funded multimodal intervention study [published correction appears in *Clin Pharmacol Ther* 1992;52:323]. *Clin Pharmacol Ther* 1992; 51: 581–585.
41. Connolly HM, Crary JL, McGoon MD, et al. Valvular heart disease associated with fenfluramine-phentermine [published correction appears in *N Engl J Med* 1997;337:1783]. *N Engl J Med* 1997; 337: 581–588.
42. Greenway FL, Whitehouse MJ, Guttadauria M, et al. Rational design of a combination medication for the treatment of obesity. *Obesity (Silver Spring)* 2009; 17: 30–39.
43. Greenway FL, Dunayvich E, Tollefson G, et al.; for the NB-201 Study Group. Comparison of combined bupropion and naltrexone therapy for obesity with monotherapy and placebo. *J Clin Endocrinol Metab* 2009; 94: 4898–4906.
44. Wadden TA, Foreyt JP, Foster GD, Hill JO, Klein S, O'Neil PM, Perri MG, Pi-Sunyer FX, Rock CL, Erickson JS, Maier HN, Kim DD, Dunayevich E. Weight Loss With Naltrexone SR/Bupropion SR Combination Therapy as an Adjunct to Behavior Modification: The COR-BMOD Trial. Obesity (Silver Spring). 2010 Jun 17. [Epub ahead of print]
45. Heymsfield SB, Greenberg AS, Fujioka K, et al. Recombinant leptin for weight loss in obese and lean adults: a randomized, controlled, dose-escalation trial. *JAMA* 1999; 282: 1568–1575.
46. Ravussin E, Smith SR, Mitchell JA, et al. Enhanced weight loss with pramlintide/metreleptin: an integrated neurohormonal approach to obesity pharmacotherapy. *Obesity (Silver Spring)* 2009; 17: 1736–1743.
47. Smith SR, Prosser WA, Donahue DJ, et al.; APD356-004 Study Group. Lorcaserin (APD356), a selective 5-HT(2C) agonist, reduces body weight in obese men and women. *Obesity (Silver Spring)* 2009; 17: 494–503.
48. Smith SR, Weissman NJ, Anderson CM, et al.; for the Behavioral Modification and Lorcaserin for Overweight and Obesity Management (BLOOM) Study Group. Multicenter, placebo-controlled trial of lorcaserin for weight management. *N Engl J Med* 2010; 363: 245–256.
49. Astrup A, Madsbad S, Breum L, et al. Effect of tesofensine on bodyweight loss, body composition, and quality of life in obese patients: a randomised, double-blind, placebo-controlled trial. *Lancet* 2008; 372: 1906–1913.

Surgical Therapy for Diabesity

Kervin Arroyo, MD, Subhash U. Kini, MD, John E. Harvey, MD
and Daniel M. Herron, MD

Mount Sinai School of Medicine, New York, NY

OUTLINE

Current Bariatric Surgical Procedures
- **Roux-en-Y Gastric Bypass**
- **Laparoscopic Adjustable Gastric Banding**
- **Biliopancreatic Diversion with Duodenal Switch**
- **Sleeve Gastrectomy**

Mechanism of Action for Resolution of Diabetes
Novel Metabolic Operations
Conclusion

ABSTRACT

Type 2 diabetes mellitus affects more than 170 million people worldwide. Because this disease is strongly linked to obesity, the term "diabesity" has been coined to describe the confluence of the 2 disease processes. Bariatric surgery has been performed for many years to achieve sustained weight loss in the morbidly obese population. As a secondary effect, a remarkable improvement in glycemic control is commonly achieved postoperatively. This has led to substantial interest in the use of bariatric procedures to treat type 2 diabetes mellitus. Surgical procedures in common use include the adjustable gastric band, the Roux-en-Y gastric bypass, the biliopancreatic diversion with duodenal switch, and the sleeve gastrectomy. Additionally, several investigational procedures including the ileal interposition and duodenal-jejunal bypass have been proposed as primary interventions for type 2 diabetes mellitus. These operations achieve their metabolic effects through a combination of volume restriction, intestinal bypass, and hormonal changes. As more data become available on the positive effect of bariatric procedures on type 2 diabetes mellitus, the use of such operations may grow. Bariatric surgery may ultimately become a major tool in the long-term treatment of type 2 diabetes mellitus. This manuscript presents an extensive review of the literature supporting these concepts. ***Mt Sinai J Med 77:418–430, 2010.*** © *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Daniel M. Herron
Department of Surgery
Mount Sinai School of Medicine
New York, NY
Email: daniel.herron@mountsinai.org

Key Words: bariatric surgery, biliopancreatic diversion, duodenal-jejunal bypass, gastric band, ileal interposition, Roux-en-Y bypass, sleeve gastrectomy, type 2 diabetes mellitus.

Type 2 diabetes mellitus (T2DM) is one of the most common diseases in the world, affecting more than 170 million people. It is estimated that by 2030, T2DM may afflict 365 million people worldwide.[1] Within the United States, there are more than 16 million diabetics; this number is expected to increase to almost 50 million by 2050.[2] T2DM accounts for 90% of all diabetes cases in the world.[3] As T2DM is strongly linked to obesity, the term "diabesity" has been coined to describe the confluence of the 2 diseases.[4]

The epidemic of T2DM is closely linked to the skyrocketing prevalence of severe obesity, currently one of the most pressing health problems facing the United States and the world. The National Health and Nutrition Examination Survey (1999–2004) estimated that roughly one-third of US adults are overweight, one-third are obese, and 5.9% percent are extremely obese. This totals 66%, or roughly two-thirds of the US population.[4,5] In 2005, the World Health Organization estimated that 1.6 billion adults were overweight and at least 400 million were obese.[6]

Numerous studies to date have suggested that diet and exercise, even when combined with weight-loss medications, do not result in durable weight loss for the severely obese patient.[7] Additionally, although it is recognized that continuous medical therapy can achieve good glycemic control in the diabetic, cessation of therapy will result in the rapid return of hyperglycemia. Additionally, some population studies suggest that more than 38% of individuals with T2DM maintain a level of glycated hemoglobin

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20209

© 2010 Mount Sinai School of Medicine

Table 1. *Bariatric Surgical Procedures and Their Mechanisms of Action.*

Procedure	Mechanism of Action
Laparoscopic adjustable gastric banding	Restriction
Sleeve gastrectomy	Restriction
Roux-en-Y gastric bypass	Restriction, malabsorption (minor), hormonal
Biliopancreatic diversion with duodenal switch	Restriction and malabsorption (major), hormonal
Ileal transposition*	Hormonal
Duodenal-jejunal bypass*	Malabsorption (minor), hormonal

* Investigational.

A_{1c} (HgA_{1c}) >7.0%. Because both sustained weight loss and durable glycemic control are commonly achieved after weight-loss surgery, there has been substantial interest in the use of bariatric procedures to treat T2DM.[5] In this article, we review the current bariatric operations, their mechanisms of action, and their ultimate effects on weight loss and control or resolution of diabetes.

CURRENT BARIATRIC SURGICAL PROCEDURES

The National Institutes of Health convened a consensus conference on bariatric surgery in 1991 that recommended surgical therapy for patients who had failed behavioral, dietary, and exercise approaches.[8] Body mass index (BMI) was the primary measure used to determine whether surgery was indicated. The consensus group stated that bariatric surgery was indicated for a BMI >40 kg/m^2, or a BMI >35 kg/m^2 in the presence of 1 or more significant weight-related comorbid diseases, such as hypertension, sleep apnea, or T2DM. As of 2010, 4 primary bariatric procedures were performed in the United States for the treatment of morbid obesity. Each operation relies on 1 or more mechanisms of weight loss, including volume restriction, malabsorption, and hormonal changes (Table 1). The operations provide good to excellent results with regard to the improvement or resolution of diabetes.

As of 2010, 4 primary bariatric procedures were performed in the United States for the treatment of morbid obesity. Each operation relies on 1 or more mechanisms of weight loss, including volume restriction, malabsorption, and hormonal changes.

Roux-en-Y Gastric Bypass

As early as 1955, it was noted that subtotal gastrectomy for peptic ulcer disease resulted in marked improvement in diabetes.[9] Based partially on these gastric resectional operations, the Roux-en-Y gastric bypass (RYGB) was first developed as a weight-loss operation in the 1960s and started being performed using minimally invasive techniques in the late 1990s (Figure 1).[10] In 2003, it was estimated that 103,000 bariatric procedures were performed in United States, 85% of which were RYGB.[11] As of 2007, RYGB represented 80% of all bariatric surgeries performed in the United States.[12] A 2009 survey by Buchwald *et al.* suggested that 112,000 RYGB operations were performed in 2008.[13] The

In 2003, it was estimated that 103,000 bariatric procedures were performed in United States, 85% of which were Roux-en-Y gastric bypass.

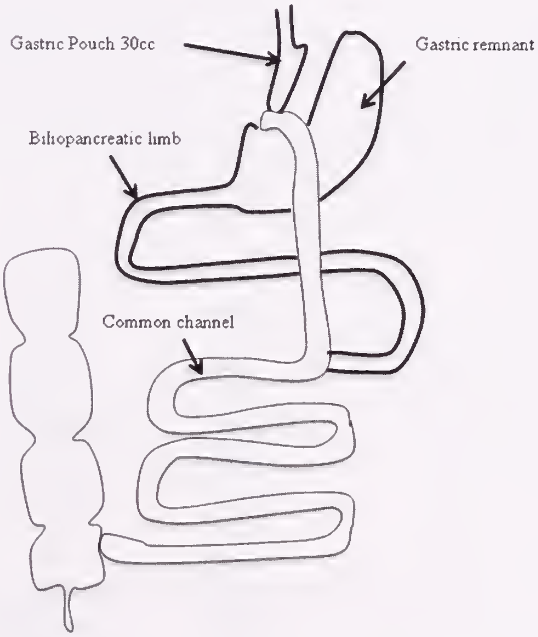


Fig 1. Roux-en-Y gastric bypass. A small gastric pouch is created, fully separated from the gastric remnant. The pouch is connected to the jejunum using a Y-shaped intestinal configuration.

mechanism of action of weight loss after RYGB has historically been described as a mix of restriction and malabsorption, but more recent investigations suggest an important hormonal aspect.[14]

In 2010, RYGB is primarily performed via a laparoscopic approach, through 5 to 7 small trocar sites. With this operation a small pouch of proximal stomach is created using a surgical stapler. The pouch volume is generally 15–30 mL. The jejunum is then divided approximately 50 to 100 cm distal to the ligament of Treitz. This limb comprises the biliopancreatic limb that transmits all the secretions from the stomach, the biliary tree, and the pancreas. The jejunum is then reconnected in a Y-shaped fashion. The distal aspect of the divided jejunum is anastomosed to the gastric pouch, and 100 to 150 cm distal to that, the proximal part of the divided jejunum is connected. The segment of bowel from the stomach pouch to the distal anastomosis is called the alimentary limb, as it contains food particles only, without bile or pancreatic juices. The bowel beyond the anastomosis is called the common channel, as it contains both food particle and digestive juices. Food passes from the esophagus to the gastric pouch, then to the alimentary limb and common channel.

Weight Loss after Roux-en-Y Gastric Bypass

RYGB results in the loss of 60% to 80% of initial excess weight (Table 2). In 1995, Dr. Walter Pories published a study of 608 patients who underwent gastric bypass, achieving an excess weight loss (EWL) of 69% at 1 year, 58% at 5 years, 55% at 10 years, and 50% at 14 years.[15] This study demonstrated the durability of long-term weight loss with the RYGB. In a more recent study by Dr. Phil Schauer's group in 2003, 1160 gastric bypass patients achieved 60% EWL at a mean follow-up period of 20 months.[16]

These statistics are echoed by many other studies. A 2003 publication from Virginia Commonwealth University described results of RYGB for 1025 patients. The authors reported an EWL of 66% at 1 year, 50% at 5 years, and 52% at 10 years.[17] MacDonald *et al.* from East Carolina University reported results from a series of 232 patients.[18] One hundred fifty-four morbidly obese patients underwent RYGB, and another 78 did not. The EWL of the RYGP group was found to be 62% at 1 year and 50% at 14 years. In 2004, Dr. Nicholas Christou's group performed a retrospective study of 802 patients who had undergone RYGB with a follow-up of 16 years. They reported EWL of 69% ± 23%.[19]

Resolution of Diabetes after Roux-en-Y Gastric Bypass

Multiple studies demonstrate the improvement or resolution of diabetes after RYGB (Table 3).

> *Multiple studies demonstrate the improvement or resolution of diabetes after Roux-en-Y gastric bypass (RYGB).*

Although many bariatric surgeons had noted an improvement in their patients' diabetes after bariatric surgery, Pories was one of the first investigators to quantify the effect. In 1995, he published a landmark article titled "Who would have thought it? An operation proves to be the most effective therapy for adult-onset diabetes mellitus."[15] The investigators reported their series of 608 RYGB

Table 2. *Percentage Excess Weight Loss After Bariatric Surgery.*

Authors	Year	No. of Patients	Procedure(s)	Follow-up	% EWL
O'Brien & Dixon[24]	2003	1025	LAGB	5 years	56
Ponce *et al.*[26]	2005	1014	LAGB	4 years	65
Sénchez-Santos *et al.*[43]	2009	549	SG	3 years	72
Cummings *et al.*[74]	2002	26	SG	5 years	55
Buchwald *et al.*[80]	2004	1848, 4204, 2480	LAGB, RYGB, BPD-DS	13 years	47, 61, 70
Arias *et al.*[46]	2009	130	SG	2 years	68
Pories *et al.*[15]	1995	680	RYGB	10 years	55
Schauer *et al.*[16]	2003	1025	RYGB	10 years	52
MacDonald *et al.*[18]	1997	154	RYGB	14 years	50
Strain *et al.*[37]	2009	49	BPD-DS	27 months	84
Rabkin *et al.*[36]	2003	345	BPD-DS	2 years	91
Ren *et al.*[34]	2000	40	BPD-DS	9 months	75
Scopinaro *et al.*[81]	2002	26	BPD	1 year	68
Buchwald *et al.*[20]	2009	135, 246	All	2 years	64

Abbreviations: BPD, biliopancreatic diversion; BPD-DS, biliopancreatic diversion with duodenal switch; EWL, excess weight loss; LAGB, laparoscopic adjustable gastric banding; RYGB, Roux-en-Y gastric bypass; SG, sleeve gastrectomy.

Table 3. *Results of Diabetes Control After Bariatric Surgery.*

Authors	Year	No. of Patients	Procedure(s)	Follow-up	Outcome: Resolution of Diabetes, %
Pories *et al.*[15]	1995	121	RYGB	14 years	83
Buchwald *et al.*[80]	2004	205, 989, 288	LAGB, RYGB, BPD-DS	13 years	48, 84, 98
Sarson *et al.*[82]	1981	191	RYGB	20 months	83
Sjöström *et al.*[21]	2004	342	RYGB, LAGB	2 years	72
Torquati *et al.*[22]	2005	117	RYGB	30 months	74
Sugerman *et al.*[56]	1989	1025	RYGB	2 years	83
Scopinaro *et al.*[38]	2008	443	BPD	20 years	91
Dixon, O'Brien[28]	2002	50	LAGB	2 years	64
Ponce *et al.*[30]	2004	53	LAGB	2 years	80
Cottam *et al.*[48]	2006	126	SG	1 year	81
Silecchia *et al.*[49]	2006	44	SG	1.9 years	76
Buchwald *et al.*[20]	2009	135,246	LAGB, RYGB, BPD-DS	2 years	57, 80, 95

Abbreviations: BPD, biliopancreatic diversion; BPD-DS, biliopancreatic diversion with duodenal switch; LAGB, laparoscopic adjustable gastric banding; RYGB, Roux-en-Y gastric bypass; SG, sleeve gastrectomy.

patients, with particular emphasis on the 121 patients with T2DM and 150 with impaired glucose tolerance, followed for up to 14 years after surgery. Of the 121 patients with T2DM, 83% maintained normal levels of plasma glucose, insulin, and HgA_{1c} without additional medical therapy. For the 150 with impaired glucose tolerance, 98.7% maintained normal levels.[15]

A more recent meta-analysis of 621 studies and 135,246 patients documented the resolution of T2DM in RYGB and laparoscopic adjustable gastric banding (LAGB). Resolution of diabetes occurred in 80.3% of RYGB patients, compared with 56.7% of those who underwent LAGB.[20] In 2003, Schauer's group from the University of Pittsburgh published results from a series of 191 RYGB patients with T2DM. After a mean follow-up of 19.7 months, 83% of RYGB patients had normal levels of fasting glucose and HgA_{1c}.[16]

The Swedish Obese Subjects Study Group represents one of the largest and best-followed bariatric surgery cohorts. In 2004, the investigators published long-term follow-up comparing surgical versus medical management. The cohort included 4047 subjects followed for 2 years and 1703 subjects followed for 10 years. Recovery from T2DM was identified in 72% of patients undergoing bariatric surgery versus 21% of those in the medical management group.[21] In another study from Vanderbilt University, Torquati *et al.* reported a series of 117 patients with 83% follow-up at 30 months. T2DM was fully resolved in 74% of the patients and improved in the remainder.[22]

> *The Swedish Obese Subjects Study Group, reported in 2004, represents one of the largest and best-followed bariatric surgery cohorts. The cohort included 4047 subjects followed for 2 years and 1703 subjects followed for 10 years. Recovery from type 2 diabetes mellitus was identified in 72% of patients undergoing bariatric surgery versus 21% of those in the medical management group.*

In the group from East Carolina University reported by MacDonald *et al.*,[18] the mean blood glucose level in the surgical group dropped from 187 mg/dL preoperatively to 99 mg/dL at 1 year and remained below 140 mg/dL for 10 years of follow-up. Additionally, the antiglycemic medication requirement fell from 31% to 9% for the RYGB group, whereas it increased to 87% in the control group at 9 years of follow-up.

A recent study from Columbia University evaluated RYGB patients with T2DM who were >3 years out from their surgery.[23] Although diabetes improved or resolved in all patients initially, it recurred in 24% of patients. Recurrence of T2DM was more likely in patients with a lower preoperative BMI and those who failed to achieve EWL of 50% or more.

Laparoscopic Adjustable Gastric Banding

The gastric band is a restrictive device consisting of a rigid plastic band with a saline-filled silicone rubber balloon lining its inner surface (Figure 2).

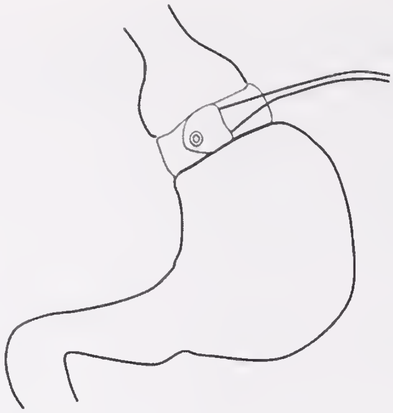


Fig 2. Laparoscopic adjustable gastric band. An adjustable band is placed around the upper aspect of the stomach, just beneath the gastroesophageal junction.

The procedure to implant the device represents an evolution of the open gastric band developed in the early 1990s. Because of the simple technical nature of the procedure, LAGB has grown substantially in popularity, becoming one of the most common bariatric procedures in 2010.

The procedure consists of placing the band around the upper stomach just beneath the gastroesophageal junction. The anterior stomach is then imbricated over the band and sutured in position, securing it in place. The band is connected, via thin flexible tubing, to a subcutaneous port that can be palpated and accessed with a hypodermic needle in the surgeon's office. Injection of saline into the system will inflate the balloon and tighten the band, while removal of saline loosens it. Usually the first adjustment is done 3 to 6 weeks after operation, with subsequent adjustments performed on a regular basis, generally every 1 to 2 months thereafter. The ultimate goal of the adjustment process is to help the patient achieve satiety after a small meal, thereby reducing ingested volume.

Weight Loss after Laparoscopic Adjustable Gastric Banding

Multiple studies have reported weight loss following LAGB. Australian investigators reported an EWL of 55% at 3 years and 56% at 5 years in a cohort of 1250 patients.[24] Weight loss was somewhat less in a New York series of 445 patients published by Ren *et al.* in 2004. At 12 months of follow-up, an EWL of 44% was reported.[25] Dr. Jaime Ponce reported his series of 1014 LAGB patients in 2005. His cohort achieved a mean EWL of 40% at 1 year and 65% at 4 years.[26]

A 2005 meta-analysis of LAGB studies reported EWL in the United States of 36% to 61% at 1 year and 56% at 3 years. Outside the United States, the EWL was similar: 38% to 56% at 1 year, 54% to 64% at 3 years, and 53% to 61% at 5 years.[27]

A 2005 meta-analysis of laparoscopic adjustable gastric bandstudies in the United States reported excess weight loss of 36% to 61% at 1 year and 56% at 3 years. Outside the United States, the excess weight loss was similar: 38% to 56% at 1 year, 54% to 64% at 3 years, and 53% to 61% at 5 years.

Resolution of Diabetes after Laparoscopic Adjustable Gastric Banding

The LAGB procedure has not been quite as successful as RYGB with respect to diabetes resolution. In an early report, Dixon *et al.* reported a study of 50 T2DM patients who underwent LAGB and were followed up after 1 year. Remission of T2DM, defined as normal HbA_{1c} without medication, was achieved in 64% of patients, whereas 26% noted improvement in glycemic control and 10% were unchanged.[28] A later randomized study by the same author demonstrated remission of T2DM in 73% of subjects treated with LAGB but in only 13% of controls treated with diet and exercise.[29] Another study performed by Ponce *et al.* retrospectively reviewed 402 LAGB patients, 53 of whom were on hypoglycemic medication for T2DM preoperatively.[30] Resolution of T2DM was noted in 66% of patients at 1 year and 80% at 2 years. Results published in 2004 reporting data from the ongoing Swedish Obesity Study revealed inferior glycemic control in LAGB as compared with RYGB.[21]

Results in 2004 from the ongoing Swedish Obesity Study revealed that, compared with RYGB, laparoscopic adjustable gastric banding (LAGB) resulted in inferior glycemic control.

Biliopancreatic Diversion with Duodenal Switch

The biliopancreatic diversion with duodenal switch is most commonly referred to by its initials, BPD-DS

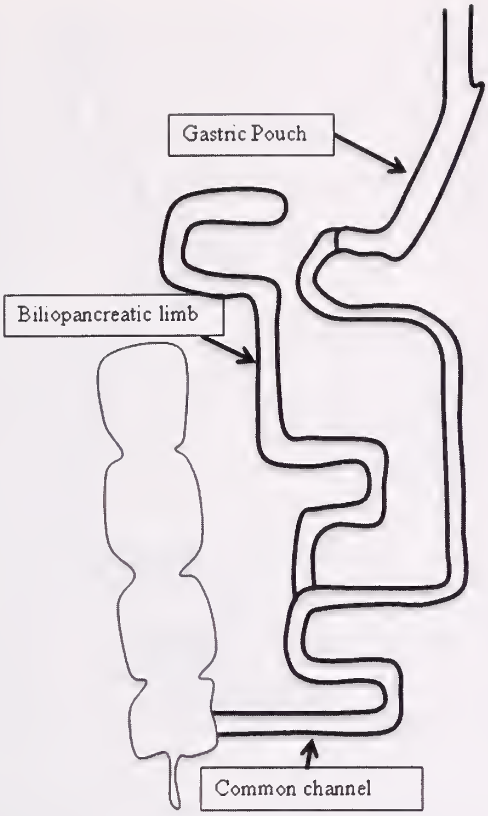


Fig 3. Biliopancreatic diversion with duodenal switch. A sleeve gastrectomy is created and the first portion of the duodenum is connected to the distal 250 cm of small bowel. The common channel is 100 cm in length. The biliopancreatic limb passes only bile and pancreatic secretions.

(Figure 3). The most technically complex bariatric operation, BPD-DS causes weight loss through a combination of restriction and substantial malabsorption. It was first described in 1988 by 2 independent groups: Marceau in Canada and Hess in Ohio.[31,32] The restriction component of the operation is achieved with a sleeve gastrectomy (SG) in which the left side, or greater curvature, of the stomach is resected, leaving a banana-shaped stomach. The duodenum is divided just past the pylorus and connected to the distalmost 250 cm of ileum. The remainder of the small intestine (second through fourth portion of duodenum, jejunum, and proximal ileum) is completely bypassed and anastomosed to the ileum roughly 100 cm from the ileocecal junction. This final segment of bowel, the "common channel," is the only portion of small intestine where food is normally mixed with bile and pancreatic enzymes. Because of the substantial malabsorption caused by such anatomy, patients often complain of frequent and soft bowel movements after this surgery. In some patients, malabsorption is so severe that parenteral nutrition and subsequent surgical revision are required. A 2003 series reported that 5% of BPD-DS patients ultimately required surgical revision for severe malnutrition.[33]

BPD-DS was initially performed using open surgical techniques through a vertical midline incision. The laparoscopic approach to the BPD-DS was first performed by Dr. Gagner at Mount Sinai Medical Center in New York in 1999, and was found to decrease postoperative pain and recovery time.[34]

Weight Loss after Biliopancreatic Diversion with Duodenal Switch

It is well-known that BPD-DS carries a higher incidence of morbidity and mortality than other bariatric procedures. However, it also demonstrates the best long-term weight loss. Scopinaro reported on 2241 patients who underwent BPD without DS, a similar operation, from 1984 through 1997.[35] Patients achieved an EWL of 74% at 2 years and 77% at 10 years. In Rabkin's 2003 series of 345 BPD-DS patients, mean EWL was 51% at 6 months, 89% at 18 months, and 91% at 2 years.[36] One of the first laparoscopic BPD-DS series was published in 2000 by the Mount Sinai group in New York; at a mean follow-up of 9 months, patients enjoyed an EWL of 75%.[34] Strain *et al.* reported 49 BPD-DS patients from New York Presbyterian Hospital who achieved an EWL of 84% at 27 months.[37]

> *It is well-known that biliopancreatic diversion with duodenal (BPD-DS) carries a higher incidence of morbidity and mortality than other bariatric procedures. However, it also demonstrates the best long-term weight loss.*

Resolution of Diabetes after Biliopancreatic Diversion with Duodenal Switch

Of all bariatric operations, BPD-DS demonstrates the best results for resolution or improvement of T2DM (Table 3). In Scopinaro's report of 2241 BPD patients, those with T2DM had 100% resolution of diabetes at 4 months after the operation.[35] In a 2008 study, Scopinaro's group compared RYGB with BPD, collecting data from 443 patients with T2DM from 1976 to May 2007.[38] Within 2 months after surgery, 75% of patients had a fasting serum glucose (FSG) level within normal limits. A more significant drop was achieved at 1 year follow-up, and FSG levels remained stable thereafter for almost 20 years. He reported that the remission of the diabetes after BPD-DS was inversely related to the duration and the treatment of diabetes before the operation.

In Buchwald's 2004 meta-analysis, he reported that T2DM resolved in 48% of LAGB patients, 84% of RYGB patients, and 98% of patients who underwent BPD-DS.[39] In a 2008 meta-analysis that included a total of 135,246 patients, overall resolution of T2DM for all bariatric procedures was 78%, with improvement or resolution in 87%.[20] When procedures were evaluated independently, the authors found resolution of T2DM in 95% of BPD-DS patients, 80% of RYGB patients, and 57% of LAGB patients. When looking at bariatric procedures as a whole, diabetes parameters improved significantly, with insulin levels dropping by 97 mU/L, HbA_{1c} decreasing by 2.1%, and FSG decreasing by 44 mmol/L.[20]

Of all bariatric operations, BPD-DS demonstrates the best results for resolution or improvement of type 2 diabetes mellitus (T2DM). In a 2008 meta-analysis involving a total of 135,246 patients, the resolution of T2DM in BPD-DS patients was 95%; in RYGB patients, 80%; and in LAGB patients, 57%.

Sleeve Gastrectomy

SG represents the most recently developed bariatric operation in common use as of 2010. The SG was initially performed as the gastric component of the BPD-DS.[40] In 2003, Gagner's group at the Mount Sinai School of Medicine reported the use of SG as the first stage of a 2-stage BPD-DS.[40] It was felt that performing the technically straightforward SG on patients with a BMI >60 kg/m^2 would allow the patients to lose a substantial amount of weight, allowing for the remainder of the BPD-DS to be completed with lower overall morbidity. After many patients were found to obtain good weight loss with the "first stage" alone, use of the SG as a primary procedure grew substantially. Conceptually, SG is a purely restrictive procedure. However, some hormonal affects may be achieved by the removal of hormonally active gastric tissue as well as by alteration of gastric emptying.

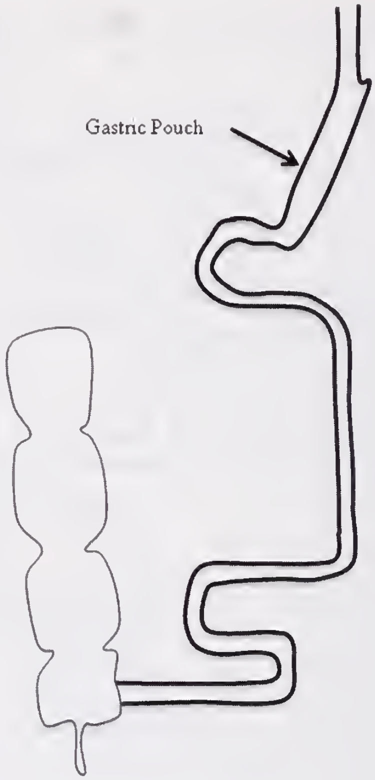


Fig 4. Sleeve gastrectomy. The greater curvature of the stomach is resected, leaving a banana-shaped stomach.

In SG (Figure 4), a 10–14 mm-diameter bougie (sizing tube) is placed transorally until it lies along the right side, or lesser curvature, of the stomach. A surgical stapler is then repeatedly fired adjacent to the tube until the greater curvature (left side) of the stomach is completely resected. This results in a small, banana-shaped stomach with a volume of roughly 150–200 mL. In addition to causing volume restriction, SG decreases ghrelin levels and may thereby decrease appetite.[41] Some studies suggest that SG may present a lower morbidity than RYGB.[42]

Weight Loss After Sleeve Gastrectomy

Generally, weight loss after SG falls between that of LAGB and RYGB. A 2009 study from the Spanish national registry presented data from 17 centers and a total of 540 patients (Table 2). Patients achieved an EWL of 56% at 6 months, 68% at 1 year, and 72% at 2 and 3 years.[43] An Australian study reported a series of 185 patients with an EWL of 47% at 1 year, 60% at 2 years, and 66% at 3 years.[44] A 2009 study from Jackson Memorial Hospital in Florida revealed an EWL of 78% at 1 year and 75% at 2 years.[45] Somewhat lower weight loss was reported in a 2009 report from the Cleveland Clinic in Florida: 130 SG patients enjoyed a mean EWL of 51% at 6 months, 62% at 1 year, and 68% at 2 years.[46] Conglomerated data representing about 14,000 patients were presented at the Second International Consensus Conference on Sleeve Gastrectomy in March 2009; EWL was reported to be 60% at 1 year, 64% at 2 years, and 62% at 4 years.[47]

Generally, weight loss after sleeve gastrectomy falls between that of laparoscopic adjustable gastric banding and Roux-en-Y gastric bypass. At the Second International Consensus Conference on Sleeve Gastrectomy in March 2009, conglomerated data involving about 14,000 patients showed excess weight loss to be 60% at 1 year, 64% at 2 years, and 62% at 4 years.

Resolution of Diabetes After Sleeve Gastrectomy

A 2006 study from the University of Pittsburgh described a cohort of 126 patients; after 1 year, 81% of patients with T2DM had resolution of their disease and another 10% had significant improvement (Table 3).[48] An Italian study of 44 patients with 1.9 years follow-up showed complete resolution of T2DM in 76% and improvement in 15%.[49] A group from Spain noted a slightly higher rate of resolution of T2DM, 85% at 1 year postoperative.[50] Several other groups support these numbers, reporting T2DM resolution rates ranging from 84% to 95% 1 year after surgery.[50–52]

In a 2006 study at the University of Pittsburgh involving 126 patients with T2DM, the data after 1 year showed complete resolution of T2DM in 81% of patients and significant improvement in another 10%.

At the Second International Consensus Conference on Sleeve Gastrectomy, the Italian Registry reported a resolution of T2DM in 80% of patients.[47] Additionally, they found that resolution of diabetes was independent of excess weight loss.

A recent study from Switzerland randomized 13 patients to RYGB and 14 to SG; after 3 months, no significant differences in insulin and glucagon-like peptide 1 (GLP-1) secretion were noted between the 2 groups.[53] This finding suggests that SG has a hormonal effect that may be independent of pure volume restriction.

MECHANISM OF ACTION FOR RESOLUTION OF DIABETES

It is likely that the improvement in T2DM observed after bariatric surgery is a result of several mechanisms of action. In addition to weight loss, these mechanisms may include diverse factors such as rearrangement of the gastrointestinal anatomy, intestinal malabsorption, alterations in beta-cell activity, and hormonal changes.[52] After bariatric surgery, both insulin sensitivity and insulin signaling in tissues increases, facilitating fatty-acid metabolism and decreasing intramuscular and intrahepatic lipids.[54,55]

It is likely that the improvement in type 2 diabetes mellitus observed after bariatric surgery is a result of several mechanisms of action. In addition to weight loss, these mechanisms may include diverse factors such as rearrangement of the gastrointestinal anatomy, intestinal malabsorption, alterations in beta-cell activity, and hormonal changes.

For many years it has been appreciated that weight loss results in improvement in T2DM. Interestingly, improvements are typically noted in the immediate postoperative period prior to significant weight loss.[15] Such immediate improvement is seen after operations involving restructuring of the gastrointestinal tract, such as RYGB and BPD-DS, but is less pronounced after purely restrictive procedures such as LAGB, where improvement in T2DM is mainly due to weight loss.[54] This concept is supported by many studies, such as Scopinaro's 1998 review showing that, 1 year after BPD, all patients enjoyed normalization of their blood sugar while still retaining a substantial amount of excess body weight.[35] As early as 1989, Sugerman noted that patients undergoing RYGB and BPD-DS achieved better control of their diabetes compared with gastric band patients.[56]

Several investigators have evaluated differences in diabetic control after weight loss surgery and medical therapy. Laferrère *et al.* compared a series of patients undergoing RYGB with a group treated with medical therapy.[57] After both groups had lost approximately 9 kg, the RYGB patients were found to have better glucose control despite identical weight loss.

These data support the concept that rearrangement of gastrointestinal anatomy is an important mechanism in weight loss and control of T2DM. Two different hypotheses have been proposed to explain these effects: the foregut hypothesis and the hindgut hypothesis. The foregut hypothesis states that overstimulation of the foregut–the stomach, duodenum, and jejunum–is an important causative factor in T2DM. Operations that bypass this area of the GI tract would necessarily result in decreased stimulation of the foregut, resulting in improved glycemic control.[58]

The hindgut hypothesis is an alternate, but not mutually exclusive, theory that suggests that increased stimulation of the distal small bowel is the responsible factor. By facilitating the presentation of intestinal contents to the ileum, incretin release–including GLP-1 and other peptide mediators–is enhanced, ultimately resulting in improvement in T2DM.[59–63] This hypothesis is supported by several animal studies. Strader *et al.* performed ileal transposition in rats, moving a segment of distal small bowel to a more proximal position in the GI tract. The rats that underwent ileal transposition were then compared with a group that underwent intestinal transection and reanastomosis without transposition. Rats with transposition were found to have an increase in insulin sensitivity compared with controls.[64] This hypothesis has found additional support from recent human studies that have evaluated several different versions of SG with ileal interposition and found improvement in HgA$_{1c}$ and serum triglyceride levels.[65]

Intestinal malabsorption is incurred by some bariatric procedures but not others.[66] A small amount of malabsorption is caused by RYGB, whereas BPD-DS results in significant fat and protein malabsorption.[32,67,68] Resolution of T2DM has been found to be better in more malabsorptive operations like BPD-DS, compared with RYGB.[32]

Intestinal malabsorption is incurred by some bariatric procedures but not others. A small amount of malabsorption is caused by RYGB, whereas BPD-DS results in significant fat and protein malabsorption. Resolution of T2DM has been found to be better in more malabsorptive operations like BPD-DS, compared with RYGB.

Multiple studies show greater hormonal changes after malabsorptive procedures compared with purely restrictive procedures.[59,60,64,69–71] These hormonal changes are related to improved glucose homeostasis after surgery. Incretins are defined as peptides secreted in response to a meal that enhance insulin secretion. These peptides form the basis of the anti-incretin hypothesis described by Rubino *et al.*[72] Rubino stated that "increase or untimely production of the anti-incretin signal could disrupt the incretin/anti-incretin homeostatic mechanism and ultimately affect the functions of a number of organs that are involved in the regulation of metabolism (such as beta cells, adipose tissue and the brain)." All these changes benefit glycemic control.

Postprandial plasma peptide YY is a hormone that is secreted by the L cells in the hindgut and is released into the circulation after a meal. The effect of this hormone is to promote a feeling of satiety.[71] GLP-1 and peptide YY hormones are also involved in decrease intestinal motility, gastric emptying, small intestine transit, and food intake.[71] Gastric inhibitory polypeptide, another hormone involved in glucose control, is secreted by K cells in the proximal small intestine and stimulates insulin release and synthesis.[52] Aside from the effect on insulin, the changes after surgery has been conflicting.[73,16] Leptin is an adipocyte-derived hormone with receptors in the hypothalamus; stimulation of this hormone causes a decrease in neuropeptide Y levels, suppressing appetite and increasing the metabolic rate.[52] However, the relationship between leptin and insulin is not well understood. Ghrelin is an appetite-stimulating hormone secreted by the fundus of the stomach that has received substantial interest over the past decade. Ghrelin levels vary predictably during the day and appear to increase during or immediately prior to mealtime. A landmark 2002 study by Cumming *et al.* demonstrated that this diurnal variation is suppressed after gastric bypass.[74] Ghrelin has been shown to have both stimulatory and inhibitory effects on insulin secretion.[75,76]

Although gut hormones appear to play an important role in both weight and glucose homeostasis, further research is needed to clarify the complex and oftentimes confusing relationships between them.

NOVEL METABOLIC OPERATIONS

In addition to the established bariatric procedures described above, a number of novel operations have been proposed that take advantage of some of the mechanisms described here. Ileal transposition (IT) and duodenal-jejunal bypass (DJB) are investigational operations that promise both weight control and T2DM resolution. (Figure 5) Data on these operations come mainly from animal studies at this time.

The IT procedure is based on the hindgut hypothesis, serving to increase the stimulation of the distal GI tract earlier in the postprandial process. In IT, a segment of distal ileum is isolated by stapling proximally and distally, but its mesenteric vascular supply is left intact. The jejunum is then surgically divided just distal to the ligament of Treitz and the previously isolated ileal segment is surgically transposed to this point.[61]

The DJB, in contrast, is based on the foregut hypothesis. Here, the proximal duodenum is surgically divided, as is the proximal jejunum (75 cm distal to the ligament of Treitz). A Y-shaped Roux limb is then created and connected to the proximal divided duodenum.[77]

In 2007 a group from Brazil performed a study comparing SG with IT versus diverting SG with IT. Fifty-two of the 60 (87%) achieved adequate glycemic control.[78] A study by Wang *et al.* from Taiwan compared DJB to IT. They reported that DJB improved diabetes in patients with a BMI <35 kg/m^2.[61] Findings included improvement in insulin sensitivity, lipid profile, and GLP-1 levels.

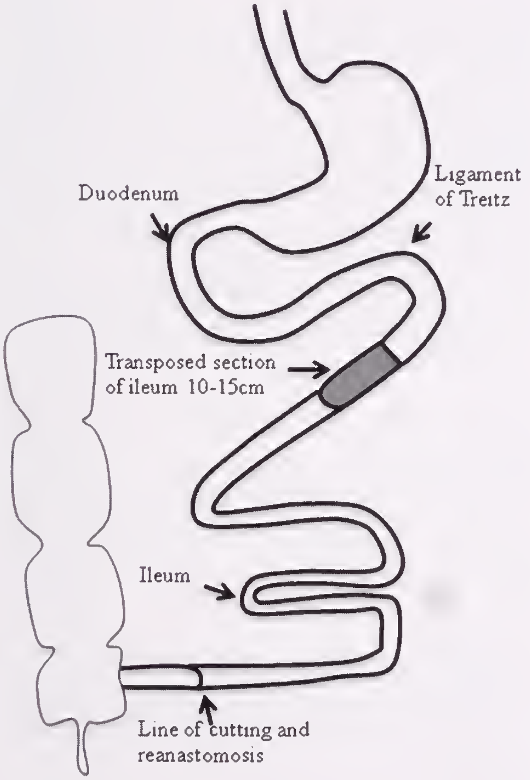

Duodenum
Ligament of Treitz
Transposed section of ileum 10-15cm
Ileum
Line of cutting and reanastomosis

Fig 5. Ileal interposition. A segment of distal ileum is isolated and reimplanted just distal to the ligament of Treitz. The divided distal ileum is reconnected.

CONCLUSION

In 2010, T2DM and obesity are increasingly significant health issues affecting the United States and many other countries throughout the world. The direct and indirect costs of the diseases and their treatment are staggering.[79] At present, bariatric surgery has been shown to durably and effectively treat both T2DM and morbid obesity.

Nonetheless, many questions regarding bariatric surgery persist. In 2010, there is no single optimal bariatric procedure, and the choice of operation must be made after a thorough discussion of risks and benefits between the surgeon and patient. Will the "perfect" bariatric procedure be identified that will serve all patients? Can patients with T2DM diabetes and BMI <35 kg/m^2 benefit from a bariatric procedure? Will different operations emerge to treat obesity versus T2DM? Will medical therapies be used jointly with surgery to provide better long-term outcomes? Answers to these questions will only come through significant future investigational efforts.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Wild S, Roglic G, Green A, et al. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 2004; 27: 1047–1053.
2. Narayan KM, Boyle JP, Geiss LS, et al. Impact of recent increase in incidence on future diabetes burden: U.S. 2005–2050. *Diabetes Care* 2006; 29: 2114–2116.

3. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. *Nature* 2001; 414: 782–787.
4. From the NIH: successful diet and exercise therapy is conducted in Vermont for "diabesity." *JAMA* 1980; 243: 519–520.
5. National Center for Health Statistics, US Department of Health and Human Services. Third National Health and Nutrition Examination Survey (NHANES III), 1988–1994, laboratory data file (CD-ROM). Hyattsville, MD: Centers for Disease Control and Prevention; 1996. Public use data file 76200. Available from National Technical Information Service (NTIS), Springfield, VA.
6. Ogden CL, Carroll MD, Curtin LR, et al. Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA* 2006; 295: 1549–1555.
7. National Heart, Lung, and Blood Institute, North American Association for the Study of Obesity. *The Practical Guide: Identification, Evaluation, and Treatment of Overweight and Obesity in Adults.* Bethesda, MD: National Institutes of Health; 2000. NIH publication 00-4084. http://www.nhlbi.nih.gov/guidelines/obesity/prctgd_c.pdf.
8. National Institutes of Health, US Department of Health and Human Services. Gastrointestinal surgery for severe obesity. Consensus Statement of the NIH Consensus Development Conference, March 25–27, 1991. http://consensus.nih.gov/1991/1991GISurgeryObesity084html.htm. Accessed 3 September 2010.
9. Friedman MN, Sancetta AJ, Magovern GJ. The amelioration of diabetes mellitus following subtotal gastrectomy. *Surg Gynecol Obstet* 1955; 100: 201–204.
10. Mason EE, Ito C. Gastric bypass in obesity: 1967 [1996 reprint]. *Obes Res* 1996; 4: 316–319.
11. Santry HP, Gillen DL, Lauderdale DS. Trends in bariatric surgical procedures. *JAMA* 2005; 294: 1909–1917.
12. Adams TD, Gress RE, Smith SC, et al. Long-term mortality after gastric bypass surgery. *N Engl J Med* 2007; 357: 753–761.
13. Buchwald H, Oien DM. Metabolic/bariatric surgery worldwide 2008. *Obes Surg* 2009; 19: 1605–1611.
14. Kashyap SR, Daud S, Kelly KR, et al. Acute effects of gastric bypass versus gastric restrictive surgery on beta-cell function and insulinotropic hormones in severely obese patients with type 2 diabetes. *Int J Obes (Lond)* 2010; 34: 462–471.
15. Pories WJ, Swanson MS, MacDonald KG, et al. Who would have thought it? An operation proves to be the most effective therapy for adult-onset diabetes mellitus. *Ann Surg* 1995; 222: 339–352.
16. Schauer PR, Burguera B, Ikramuddin S, et al. Effect of laparoscopic Roux-en Y gastric bypass on diabetes mellitus. *Ann Surg* 2003; 238: 467–484.
17. Sugerman HJ, Wolfe LG, Sica DA, Clore JN. Diabetes and hypertension in severe obesity and effects of gastric bypass-induced weight loss. *Ann Surg* 2003; 237: 751–756.
18. MacDonald KG Jr, Long SD, Swanson MS, et al. The gastric bypass operation reduces the progression and mortality of non-insulin-dependent diabetes mellitus. *J Gastrointest Surg* 1997; 1: 213–220.
19. Christou NV, Sampalis JS, Liberman M, et al. Surgery decreases long-term mortality, morbidity, and health care use in morbidly obese patients. *Ann Surg* 2004; 240: 416–424.
20. Buchwald H, Estok R, Fahrbach K, et al. Weight and type 2 diabetes after bariatric surgery: systematic review and meta-analysis. *Am J Med* 2009; 122: 248–256.
21. Sjöström L, Lindroos AK, Peltonen M, et al. Lifestyle, diabetes and cardiovascular risk factors 10 years after bariatric surgery. *N Engl J Med* 2004; 351: 2683–2693.
22. Torquati A, Lutfi R, Abumrad N, Richards WO. Is Roux-en-Y gastric bypass surgery the most effective treatment for type 2 diabetes mellitus in morbidly obese patients? *J Gastrointest Surg* 2005; 9: 1112–1118.
23. DiGiorgi M, Rosen DJ, Choi JJ, et al. Re-emergence of diabetes after gastric bypass inpatients with mid- to long-term follow-up. *Surg Obes Relat Dis* 2010; 6: 249–253.
24. O'Brien PE, Dixon JB. Lap-band: outcomes and results. *J Laparoendosc Adv Surg Tech A* 2003; 13: 265–270.
25. Ren CJ, Weiner M, Allen JW. Favorable early results of gastric banding for morbid obesity: the American experience. *Surg Endosc* 2004; 18: 543–546.
26. Ponce J, Paynter S, Fromm R. Laparoscopic adjustable gastric banding: 1,014 consecutive cases. *J Am Coll Surg* 2005; 201: 529–535.
27. Ponce J, Dixon JB. Laparoscopic adjustable gastric banding. *Surg Obes Relat Dis* 2005; 1: 310–316.
28. Dixon JB, O'Brien PE. Health outcomes of severely obese type 2 diabetic subjects 1 year after laparoscopic adjustable gastric banding. *Diabetes Care* 2002; 25: 358–363.
29. Dixon JB, O'Brien PE, Playfair J, et al. Adjustable gastric banding and conventional therapy for type 2 diabetes. *JAMA* 2008; 299: 316–323.
30. Ponce J, Haynes B, Paynter S, et al. Effect of Lap-Band-induced weight loss on type 2 diabetes mellitus and hypertension. *Obes Surg* 2004; 14: 1335–1342.
31. Marceau P, Hould FS, Simard S, et al. Biliopancreatic diversion with duodenal switch. *World J Surg* 1998; 22: 947–954.
32. Hess DS, Hess DW. Biliopancreatic diversion with a duodenal switch. *Obes Surg* 1998; 8: 267–282.
33. Anthone GJ, Lord RV, DeMeester TR, et al. The duodenal switch operation for the treatment of morbid obesity. *Ann Surg* 2003; 238: 618–627.
34. Ren CJ, Patterson E, Gagner M. Early results of laparoscopic biliopancreatic diversion with duodenal switch: a case series of 40 consecutive patients. *Obes Surg* 2000; 10: 514–523.
35. Scopinaro N, Adami GF, Marinari GM, et al. Biliopancreatic diversion. *World J Surg* 1998; 22: 936–946.
36. Rabkin RA, Rabkin JM, Metcalf B, et al. Laparoscopic technique for performing duodenal switch with gastric reduction. *Obes Surg* 2003; 13: 263–268.
37. Strain GW, Gagner M, Pomp A, et al. Comparison of weight loss and body composition changes with four surgical procedures. *Surg Obes Relat Dis* 2009; 5: 582–587.
38. Scopinaro N, Papadia F, Camerini G, et al. A comparison of a personal series of biliopancreatic diversion and literature data on gastric bypass help to explain the mechanisms of resolution of type 2 diabetes by the two operations. *Obes Surg* 2008; 18: 1035–1038.
39. Buchwald H, Avidor Y, Braunwald E, et al. Bariatric surgery: a systematic review and meta-analysis. *JAMA*. 2004; 292: 1724–1737.

40. Regan JP, Inabnet WB, Gagner M, Pomp A. Early experience with two-stage laparoscopic Roux-en-Y gastric bypass as an alternative in the super-super obese patient. *Obes Surg* 2003; 13: 861–864.
41. Bohdjalian A, Langer FB, Shakeri-Leidenmühler S, et al. Sleeve gastrectomy as sole and definitive bariatric procedure: 5-year results for weight loss and ghrelin. *Obes Surg* 2010; 20: 535–540.
42. Sammour T, Hill AG, Singh P, et al. Laparoscopic sleeve gastrectomy as a single-stage bariatric procedure. *Obes Surg* 2010; 20: 271–275.
43. Sénchez-Santos R, Masdevall C, Baltasar A, et al. Short- and mid-term outcomes of sleeve gastrectomy for morbid obesity: the experience of the Spanish National Registry. *Obes Surg* 2009; 19: 1203–1210.
44. Armstrong J, O'Malley SP. Outcomes of sleeve gastrectomy for morbid obesity: a safe and effective procedure? *Int J Surg* 2010; 8: 69–71.
45. Jacobs M, Bisland W, Gomez E, et al. Laparoscopic sleeve gastrectomy: a retrospective review of 1-and 2-year results. *Surg Endosc* 2010; 24: 781–785.
46. Arias E, Martínez PR, Ka Ming Li V, et al. Mid-term follow-up after sleeve gastrectomy as a final approach for morbid obesity. *Obes Surg* 2009; 19: 544–548.
47. Gagner M, Deitel M, Kalberer TL, et al. The Second International Consensus Summit for Sleeve Gastrectomy, March 19–21, 2009. *Surg Obes Relat Dis* 2009; 5: 476–485.
48. Cottam D, Qureshi FG, Mattar SG, et al. Laparoscopic sleeve gastrectomy as an initial weight-loss procedure for high-risk patients with morbid obesity. *Surg Endosc* 2006; 20: 859–886.
49. Silecchia G, Boru C, Pecchia A, et al. Effectiveness of laparoscopic sleeve gastrectomy (first stage of biliopancreatic diversion with duodenal switch) on co-morbidities in super-obese high-risk patients. *Obes Surg* 2006; 16: 1138–1144.
50. Vidal J, Ibarzabal A, Romero F, et al. Type 2 diabetes mellitus and the metabolic syndrome following sleeve gastrectomy in severely obese subjects. *Obes Surg* 2008; 18: 1077–1082.
51. Gan SS, Talbot ML, Jorgensen JO. Efficacy of surgery in the management of obesity-related type 2 diabetes mellitus. *ANZ J Surg* 2007; 77: 958–962.
52. Herron DM, Tong W. Role of surgery in management of type 2 diabetes mellitus. *Mt Sinai J Med* 2009; 76: 281–293.
53. Peterli R, Wölnerhanssen B, Peters T, et al. Improvement in glucose metabolism after bariatric surgery: comparison of laparoscopic Roux-en-Y gastric bypass and laparoscopic sleeve gastrectomy. *Ann Surg* 2009; 250: 234–241.
54. Rubino F, Schauer PR, Kaplan LM, et al. Metabolic surgery to treat type 2 diabetes: clinical outcomes and mechanisms of action. *Annu Rev Med* 2010; 61: 393–411.
55. Thaler JP, Cummings DE. Hormonal and metabolic mechanisms of diabetes remission after gastrointestinal surgery. *Endocrinology* 2009; 150: 2518–2525.
56. Sugerman HJ, Londrey GL, Kellum JM, et al. Weight loss with vertical banded gastroplasty and Roux-Y gastric bypass for morbid obesity with selective versus random assignment. *Am J Surg* 1989; 157: 93–102.
57. Laferrère B, Teixeira J, McGinty J, et al. Effect of weight loss by gastric bypass surgery versus hypocaloric diet on glucose and incretin levels in patients with type 2 diabetes. *J Clin Endocrinol Metab* 2008; 93: 2479–2485.
58. Rubino F, Forgione A, Cummings DE, et al. The mechanism of diabetes control after gastrointestinal bypass surgery reveals a role of the proximal small intestine in the pathophysiology of type 2 diabetes. *Ann Surg* 2006; 244: 741–749.
59. Mason EE. Ileal [correction of ilial] transposition and enteroglucagon/GLP-1 in obesity (and diabetic?) surgery. *Obes Surg* 1999; 9: 223–228.
60. Mason EE. The mechanisms of surgical treatment of type 2 diabetes mellitus. *Obes Surg* 2005; 15: 459–461.
61. Wang TT, Hu SY, Gao HD, et al. Ileal transposition controls diabetes as well as modified duodenal jejunal bypass with better lipid lowering in a nonobese rat model of type II diabetes by increasing GLP-1. *Ann Surg* 2008; 247: 968–975.
62. Frezza EE. Are we closer in finding the treatment for type 2 diabetes mellitus in morbid obesity? Are the incretins the key to success? *Obes Surg* 2004; 14: 999–1005.
63. Rubino F, Zizzari P, Tomasetto C, et al. The role of the small bowel in the regulation of circulating ghrelin levels and food intake in the obese Zucker rat. *Endocrinology* 2005; 146: 1745–1751.
64. Strader AD, Vahl TP, Jandacek RJ, et al. Weight loss through ileal transposition is accompanied by increased ileal hormone secretion and synthesis in rats. *Am J Physiol Endocrinol Metab* 2005; 288: E447–E453.
65. De Paula AL, Stival AR, Macedo A, et al. Prospective randomized controlled trial comparing 2 versions of laparoscopic ileal interposition associated with sleeve gastrectomy for patients with type 2 diabetes and BMI 21–34 kg/m^2. *Surg Obes Relat Dis* 2010; 6: 296–304.
66. Brolin RE, LaMarca LB, Kenler HA, Cody RP. Malabsorptive gastric bypass in patients with superobesity. *J Gastrointest Surg* 2002; 6: 195–205.
67. Hess DS, Hess DW, Oakley RS. The biliopancreatic diversion with the duodenal switch: results beyond 10 years. *Obes Surg* 2005; 15: 408–416.
68. Sugerman HJ. Bariatric surgery for severe obesity. *J Assoc Acad Minor Phys* 2001; 12: 129–136.
69. Korner J, Bessler M, Cirilo LJ, et al. Effects of Roux-en-Y gastric bypass surgery on fasting and postprandial concentrations of plasma ghrelin, peptide YY, and insulin. *J Clin Endocrinol Metab* 2005; 90: 359–365.
70. Farilla L, Bulotta A, Hirshberg B, et al. Glucagon-like peptide 1 inhibits cell apoptosis and improves glucose responsiveness of freshly isolated human islets. *Endocrinology* 2003; 144: 5149–5158.
71. Kini S, Herron DM, Yanagisawa RT. Bariatric surgery for morbid obesity: a cure for metabolic syndrome? *Med Clin North Am* 2007; 91: 1255–1271.
72. Rubino F, R'bibo SL, del Genio F, et al. Metabolic surgery: the role of the gastrointestinal tract in diabetes mellitus. *Nat Rev Endocrinol* 2010; 6: 102–109.
73. Kellum JM, Kuemmerle JF, O'Dorisio TM, et al. Gastrointestinal hormone responses to meals before and after gastric bypass and vertical banded gastroplasty. *Ann Surg* 1990; 211: 763–770.
74. Cummings DE, Weigle DS, Frayo RS, et al. Plasma ghrelin levels after diet-induced weight loss or gastric bypass surgery. *N Engl J Med* 2002; 346: 1623–1630.

75. Reynolds CB, Elias AN, Whisnant CS. Effects of feeding pattern on ghrelin and insulin secretion in pigs. *Domest Anim Endocrinol* 2010; 39: 90–96.
76. Gasco V, Beccuti G, Marotta F, et al. Endocrine and metabolic actions of ghrelin. *Endocr Dev* 2010; 17: 86–95.
77. Ferzli GS, Dominique E, Ciaglia M, et al. Clinical improvement after duodenojejunal bypass for nonobese type 2 diabetes despite minimal improvement in glycemic homeostasis. *World J Surg* 2009; 33: 972–979.
78. De Paula AL, Macedo AL, Rassi N, et al. Laparoscopic treatment of metabolic syndrome in patients with type 2 diabetes mellitus. *Surg Endosc* 2008; 22: 2670–2678.
79. Withrow D, Alter DA. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. *Obes Rev* 2010;doi:10.1111/j.1467-789X.2009.00712.x.
80. Buchwald H, Avidor Y, Braunwald E, et al. Bariatric surgery: a systematic review and meta-analysis. *JAMA* 2004; 292: 1724–1737.
81. Scopinaro N, Marinari GM, Camerini G. Laparoscopic standard biliopancreatic diversion: technique and preliminary results. *Obes Surg* 2002; 12: 362–365.
82. Sarson DL, Scopinaro N, Bloom SR. Gut hormone changes after jejunoileal (JIB) or biliopancreatic (BPB) bypass surgery for morbid obesity. *Int J Obes* 1981; 5: 471–480.

Metabolic Management following Bariatric Surgery

Erika Strohmayer, MD,[1] Michael A. Via, MD,[2] and Robert Yanagisawa, MD[1]

[1]Mount Sinai School of Medicine, New York, NY
[2]Albert Einstein College of Medicine, Beth Israel Medical Center, New York, NY

OUTLINE

Postoperative Nutrition
Protein-Calorie Malnutrition
Micronutrient Deficiencies
Bone Metabolism
Nephrolithiasis
Diabetes
Polycystic Ovarian Syndrome and Female Fertility
Male Sex Hormones and Fertility
Summary

ABSTRACT

Bariatric surgery is an effective treatment option for obesity. Commonly utilized procedures are either restrictive, malabsorptive, or both. Substantial weight loss can be achieved. Postoperatively, patients experience nutritional, metabolic, and hormonal changes that have important clinical implications. The postoperative diet should be advanced carefully, according to protocol. Micronutrient deficiencies such as vitamin C, vitamin A, and zinc deficiencies are common, especially following malabsorptive procedures. Bone metabolism is greatly affected, in part due to vitamin D deficiency, decreased calcium absorption, and secondary hyperparathyroidism. Diabetes improves acutely in malabsorptive procedures and in sequence with weight loss in restrictive procedures. Polycystic ovarian syndrome improves in nearly all women with this condition who undergo bariatric surgery. Testosterone levels in men also improve after surgery. Consideration of these nutritional, metabolic, and hormonal changes allows for optimal medical management following bariatric surgery. ***Mt Sinai J Med 77:431–445, 2010.*** © *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Erika Strohmayer
Mount Sinai School of Medicine
New York, NY
Email: erika.strohmayer@
mountsinai.org

Key Words: bariatric surgery, laparoscopic band surgery, malabsorptive bariatic surgery, metabolic management, postoperative malnutrition, Roux-en-Y gastric bypass.

Obesity, defined by a body mass index (BMI) of ≥ 30 kg/m^2, has reached epidemic proportions in the United States and worldwide.[1] Bariatric or weight-loss surgery is an effective tool for the treatment of obesity.[2,3] Compared with conventional diet and weight-loss strategies, a higher degree of sustained weight loss, resulting in decreased mortality and morbidity, can be obtained with bariatric surgery.[4] By current recommendations, patients with class 3 obesity (BMI ≥ 40 kg/m^2), or class 2 obesity (BMI ≥ 35 kg/m^2) with at least 1 condition associated with obesity, such as uncontrolled type 2 diabetes mellitus (T2DM) or obstructive sleep apnea, may be considered for bariatric surgery.[5]

Several bariatric procedures are presently available that utilize either gastric restriction, intestinal malabsorption, or both as strategies for the induction of weight loss. These are described in greater detail elsewhere in this issue. In brief, gastric restrictive procedures produce weight loss by limiting food intake and causing early satiety, while malabsorptive procedures directly reduce dietary caloric delivery by limiting intestinal contact with digested materials.

Vertical banded gastroplasty (VBG), commonly known as "stomach stapling," was popular in the 1980s and is one example of a gastric restrictive procedure. Nowadays, laparoscopic adjustable gastric banding (LAGB) has replaced VBG as the most common restrictive procedure.[6] Sleeve gastrectomy (SG) is also restrictive and is growing in popularity.[7] The most common bariatric procedure performed in the United States to date is the Roux-en-Y

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20211

© 2010 Mount Sinai School of Medicine

gastric bypass (RYGB), which is both restrictive and malabsorptive.[6] The biliopancreatic diversion with or without duodenal switch (BPD-DS) is highly malabsorptive and its use is declining (Figure 1).[7]

The different bariatric procedures produce variable degrees of weight loss, measured either by absolute weight loss or percentage of excess body weight lost (EBWL). Overall, malabsorptive procedures result in more weight loss than restrictive procedures. A recent meta-analysis reported mean absolute weight loss of 32 kg for gastric banding (GB), 36 kg for VBG, 45 kg for RYGB, and 44 kg for BPD-DS.[8] Percentage EBWL was reported as 46% for GB, 56% for VBG, 60% for RYGB, and 64% for BPD-DS. SG was not included in this analysis, but reports of 33%–66% EBWL are available in the literature.[9]

Through massive weight loss and possibly weight-unrelated mechanisms, bariatric surgery dramatically improves obesity and obesity-associated conditions. Endocrine disorders such as T2DM, polycystic ovarian syndrome (PCOS), and infertility can be affected. Alterations in calcium homeostasis and bone metabolism can also occur. The hormonal, nutritional, and metabolic changes that occur in the postoperative period should be appropriately managed. The purpose of this review is to elucidate metabolic and endocrine factors that should be considered in the care of the bariatric surgery patient.

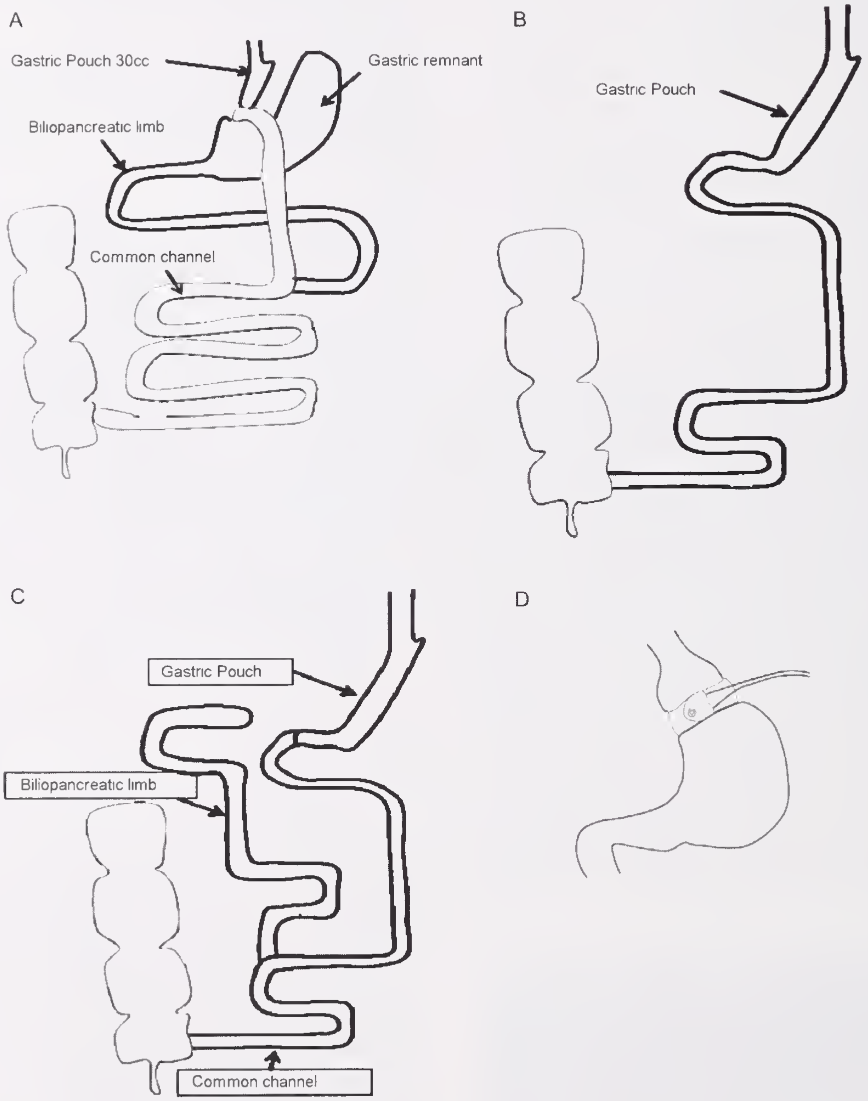


Fig 1. Schematic representation of common bariatric procedures. (A) Roux-en-Y gastric bypass, (B) sleeve gastrectomy, (C) biliopancreatic diversion with duodenal switch, (D) laparoscopic adjustable gastric banding. Courtesy D.M. Herron, MD, New York, NY.

POSTOPERATIVE NUTRITION

Dietary instructions are an important part of postoperative management. Many surgeons use dietary protocols to guide patients. An example adapted from the recent American Association of Clinical Endocrinologists perioperative bariatric surgery guidelines is provided in Table 1.[9] These protocols involve

Table 1. *Diet Stages Following Bariatric Surgery*[*].

Diet Stage	Postoperative Period	Fluids/Foods	Comments
1	1–2 days	Clear liquids[†]	Sugar-free, caffeine-free, noncarbonated
2	2–3 days	Clear and full liquids	Full liquids should contain ≤15 g sugar/serving and ≤3 g fat/serving. Examples: 1% or skim milk mixed with whey or soy protein powder (limit 20 g protein/serving) Lactaid milk or soy milk mixed with soy protein powder Light or plain yogurt Begin supplementation: chewable multivitamin w/mineral supplementation, chewable or liquid calcium citrate, vitamin D, vitamin B12[‡]
3	10–14 days	Clear liquids; replace full liquids with puréed, diced, soft, moist protein sources	Examples of protein sources: eggs, ground meat, poultry, fish, cooked beans, cottage cheese, low-fat yogurt. Nonfat gravy or light mayonnaise can be added to moisten. Continue to drink 48–64 oz. clear liquids per day, at least 30 minutes separate from food.
	4 weeks[§]	Add well-cooked soft vegetables and/or soft, peeled fruit	Protein should be consumed first to ensure adequate protein intake before satiety. Continue to drink 48–64 oz. fluid per day, at least 30 minutes separate from food.
3	5 weeks[§]	Add crackers with protein	Continue to consume protein with some fruit and/or vegetable at every meal. Avoid pasta, rice, bread. Continue to drink 48–64 oz. fluid per day, at least 30 minutes separate from food.
4	6 weeks[§]	Healthy solid food	Mixed meals of protein, fruits, vegetables, and whole-grain carbohydrates. Advise eating from small plates with small utensils for portion control. Continue to drink 48–64 oz. fluid per day, at least 30 minutes separate from food.

[*] Diet stages are not standardized. Timing for diet advancement should be modified based on the individual's tolerance for the previous diet stage.[13]

[†] For procedures other than laparoscopic adjustable gastric banding, diet can be started after a Gastrografin swallow test is performed to assess for leaks.

[‡] For biliopancreatic diversion-duodenal switch (BPD-DS), at least 350–500 μg oral vitamin B12 daily is recommended; may need to be intramuscular.

[§] Usually longer for BPD-DS.

different stages of diet advancement, starting with liquid diets and progressing through puréed and eventually solid foods over a period of weeks, depending on patient tolerance and the type of bariatric surgery. Typically sugar-free clear liquids can be started 24 hours after surgery, or sooner after LAGB. After 2 or 3 days, low-sugar, protein-rich full liquids (eg, yogurt, low-fat milk) can be added and patients should be encouraged to drink a total of 6 to 8 eight-ounce servings per day. Lactose intolerance can occur after surgery due to decreased lactase availability with malabsorptive procedures.[10] Lactose-free milk can be tried as an alternative.

Over the next 2 to 4 weeks, full liquids are replaced by puréed protein-rich foods such as eggs, poultry, and fish. Red meat is generally not tolerated until after 4 weeks.[10] Once consistently eating puréed protein, well-cooked vegetables and soft fruit can be added. If this diet is sustained, patients can start eating solid, healthy, balanced meals after about 6 weeks. An average of 60–120 g of protein/day is recommended. Protein should be consumed before other parts of mixed meals to avoid protein deficiencies that may result if satiety is reached before adequate protein consumption. Sufficient liquid intake should be maintained to avoid dehydration through all diet stages.

Gradual dietary advancement is recommended to prepare patients for the restricted diet they will maintain over the long term. Slow meal progression minimizes the likelihood of repeated vomiting that can cause anastomotic leaks and further complications.[10] There is controversy over recommended meal frequency. One study showed consumption of more frequent meals is associated with less weight loss 9 months after RYGB surgery.[11] However, many argue that small, frequent meals are beneficial because they can decrease discomfort, vomiting, and dumping syndrome (Table 1).[10,12,13]

Dumping syndrome is a common complication of RYGB surgery, with a reported prevalence of 76% in one series.[14] Symptoms include nausea, vomiting, abdominal pain, diarrhea, flushing, diaphoresis, weakness, and dizziness after intake of highly concentrated liquid or solid food. These result from rapid gastric emptying causing hypertonic chyme to enter the small intestine, which activates stretch receptors to stimulate satiety, nausea, and pain. The hyperosmolar load also induces the release of vasoactive agents and exerts large fluid shifts into the intestinal lumen, which may lead to decreased intravascular volume, low blood pressure, and tachycardia.[15] Many consider dumping syndrome a beneficial consequence of RYGB because the consequential symptoms can lead to an aversion of high-caloric foods that cause weight gain.[10] On the other hand, the association of weight loss and dumping syndrome has not been well-established, and the symptoms can be greatly debilitating.

Dumping syndrome can be prevented by adhering to specific diet and behavioral modifications. Simple carbohydrates should be avoided. Liquids should be ingested at least 30 minutes after a meal. Increased fiber intake may slow gastric emptying. Patients should chew food thoroughly and consume multiple small meals. Lying down after a meal decreases gastric emptying and helps prevent or minimize symptoms of dizziness if they occur. In severe cases that are unresponsive to these changes, medical therapy with somatostatin analogs, or surgical revision or reversal may be required.[15]

Major dietary adjustments are required for safe, successful weight loss following bariatric surgery, and the assessment of a patient's ability to make these changes should be done before surgery.[9]

PROTEIN-CALORIE MALNUTRITION

It is the intention of bariatric surgery to limit dietary intake and restrict the calories delivered to an individual, which can result in significant weight loss and improvements in obesity-associated conditions.[2,3] In some cases, bariatric surgery causes excessive caloric deficits leading to protein-calorie malnutrition.[16,17] This extreme condition may be seen following malabsorptive procedures at rates of approximately 1%–5% in RYGB patients and 4%–18% in BPD patients.[18–22] Restrictive procedures generally do not cause protein-calorie malnutrition.[17]

Symptoms of protein-calorie malnutrition can manifest as excessively rapid weight loss, muscle wasting, edema, and low serum albumin levels.[17] These are generally observed 1–2 years following RYGB and within 1 year of BPD surgery.[18] Some of the contributing factors to protein-calorie malnutrition in this population include an aversion to protein-rich foods that commonly develops postoperatively, decreased caloric intake, and decreased amino acid absorption.[16] Approximately 1% of malabsorptive surgery patients are hospitalized annually for protein-calorie malnutrition and they may require parenteral nutrition for 3–4 weeks or longer.[23] Patients who undergo malabsorptive procedures should be screened every 6 months to 1 year for protein-calorie malnutrition. Assessment should include markers of protein stores such as serum albumin or prealbumin levels.[13]

MICRONUTRIENT DEFICIENCIES

Deficiencies of micronutrients are common following both restrictive and malabsorptive weight-loss surgeries, with higher rates of both vitamin and mineral deficiencies reported in the latter.[16,24,25] Micronutrient deficiencies may result from decreased intestinal surface area in malabsorptive procedures as well as decreased micronutrient bioavailability due to decreased gastric acid production and changes in food preference observed in restrictive procedures.[16] Even before surgery, a high prevalence of micronutrient deficiencies has been observed in the obese population.[26–28] In this section we review the common micronutrient deficiencies before and after bariatric surgery with suggestions for screening and supplementation.

In general, the western diet contains food items that are nutritionally unbalanced, highly processed, and of poor nutritional value.[29] Despite an abundance of fat, salt, carbohydrates, and calories in the diet of obese patients, a paucity of micronutrients is common.[27,28] Obesity may be considered to be a "malnourished" state due to the high prevalence of vitamin and mineral deficiencies. Some authors argue that these deficiencies partially cause high caloric consumption via increased appetite stimulation in an attempt to increase intake of these micronutrients.[30]

> *Obesity may be considered to be a "malnourished" state due to the high prevalence of vitamin and mineral deficiencies.*

Common micronutrient deficiencies in the obese population before and after bariatric surgery are listed in Table 2. Before surgery, there is a high prevalence of thiamine, vitamin C, zinc, and selenium deficiencies, as well as vitamin D insufficiency.[26–28,31–33] BPD surgery has the highest rates of postoperative micronutrient deficiencies, followed by RYGB. Restrictive procedures tend to have lower rates of micronutrient deficiency (Table 2).

Regardless of the type of planned surgery, all patients considering bariatric surgery should undergo a preoperative nutritional screen that includes measurements of iron, vitamin C, vitamin B12, zinc, and vitamin D. Supplementation should be prescribed as necessary to normalize deficient levels. All patients should be given daily adult multivitamin supplements that are continued after bariatric surgery.

Postoperatively, 28%–59% of all types of bariatric surgical patients report symptoms of micronutrient deficiency, including hair loss, dry skin, muscle pain, and fatigue, within the first year after surgery.[34] Postoperative follow-up should include periodic screening for micronutrient deficiency.[13] In general, patients should be screened for micronutrient deficiencies every 6 months for the first year, and then annually. For patients who are nutritionally deficient preoperatively, initial postoperative screening should be at 2-month to 3-month intervals.[13]

Following surgery, vitamin A deficiency tends to develop in malabsorptive procedures whereas vitamin B12 deficiency occurs in both classes of bariatric surgical procedures, possibly due to decreased intrinsic factor and decreased gastric acid production.[17] Thiamine levels seem to improve postoperatively, likely due to aggressive supplementation.[16] Screening

Table 2. *Prevalence of Postoperative Micronutrient Deficiencies*.*

Vitamin/Nutrient	Preoperative Obese	SG	LAGB	RYGB	BPD-DS
Thiamine (B1)	0%–29%	0%	0%	12%	0%
Pyridoxine (B6)	0%–11%	0%	14%	10%	NA
Folate	3%–4%	22%	10%	0%–12%	5%
Cobalbumin (B12)	3%–8%	18%	0%–19%	33%–58%	22%
Vitamin C	43%	NA	48%	10%–50%	NA
Vitamin A	17%	NA	10%–25%	10%–52%	61%–69%
Vitamin D (<30 ng/dL)	80%–90%	32%	30%–40%	30%–50%	40%–100%
Vitamin E	0%	NA	0%–11%	0%–22%	4%–5%
Vitamin K	NA	NA	NA	NA	68%
Iron	26%	14%	0%–32%	25%–50%	21%–26%
Zinc	14%–30%	34%	NA	37%	10%–50%
Selenium	58%	NA	NA	3%	14.5%
Copper	NA	NA	NA	NA	70%

Abbreviations: BPD-DS, biliopancreatic diversion-duodenal switch; LAGB, laparoscopic adjustable gastric banding; NA, data not available; RYGB, Roux-en-Y gastric bypass; SG, sleeve gastrectomy.

*Table based on available reported data.[16,24,27,28,34,35,122,123] for or treating vitamin E deficiency is generally not recommended due to lack of a definitive deficiency syndrome or benefit with supplementation in adults.[13]

Patients who undergo restrictive procedures may only require basic screening that includes measurement of iron, vitamin B12, folate, vitamin D, and possibly vitamin A levels. Patients who undergo RYGB surgery should be screened for these same markers, with consideration for further testing if any deficiencies are noted.[13] BPD surgical patients should be screened for the above markers as well as for deficiencies of fat-soluble vitamins (A, D, and K), zinc, selenium, and copper.[13,35]

In addition to multivitamin preparations, supplementation to target specific deficiencies is recommended.[13] Oral vitamin K supplementation should be recommended in patients with international normalized ratio >1.4.[36] Patients who are vitamin B12 deficient require at least 350 µg oral supplementation daily.[37] Intramuscular B12 injections at 1-month to 3-month intervals may also be used to normalize levels.[38] Current recommendations for vitamin D supplementation range from 800 to as high as 150,000 IU vitamin D daily, depending on the initial level and the degree of malabsorption present.[13,39] Iron, zinc, folate, vitamin C, vitamin A, and selenium deficiencies should be treated with supplementation of the individual micronutrient that is deficient, and levels should be closely monitored.[13] Aggressive postoperative micronutrient supplementation is required to maintain adequate nutrition in this high-risk patient population.

BONE METABOLISM

The metabolism of bone is greatly affected by malnutrition and weight loss associated with bariatric surgery.[40–44] Rapid and sustained loss of bone mineralization has been documented following malabsorptive procedures.[40,41,43,44] Bone mineral density declines by approximately 10% at the hip and 5% at the spine in the first year following RYGB surgery.[40,43,45]

> *Bone mineral density declines by approximately 10% at the hip and 5% at the spine in the first year following Roux-en-Y gastric bypass surgery.*

The results of 3 small studies demonstrate that restrictive procedures have a minimal impact on bone metabolism compared with malabsorptive procedures.[41,46,47] Patients who undergo restrictive procedures appear to have stable bone mineral density and experience a decline in alkaline phosphatase activity, despite similar baseline nutritional deficits that include vitamin D insufficiency and high rates of secondary hyperparathyroidism.[32,33,41,46,47]

Patients who undergo malabsorptive procedures also experience a high degree of bone resorption that may be partly responsible for declines in bone mineral density.[40–42] At 1 year after RYGB surgery, urinary N-telopeptide levels increased 2-fold, serum C-telopeptide levels increased by 30%, and serum osteocalcin levels increased by 40% compared with baseline, indicating both increased resorption and deposition.[45,48] In a study that compared RYGB patients to obese controls, urinary N-telopeptide levels were elevated 3–4-fold 3 months after surgery, which persisted for 9 months of follow-up.[40]

High levels of osteoclast activity following RYGB surgery as evidenced by elevated C-telopeptide and N-telopeptide have led some clinicians to suggest use of antiresorptive agents such as bisphosphonate therapy to prevent bone loss in these patients.[13] Presently, there are no clinical trials that investigate the use of bisphosphonates in this population. For bariatric surgical patients that otherwise meet requirements for bisphosphonate therapy (eg, osteoporosis; bone mineral density T score <−2.5), intravenous bisphosphonates may be preferred due to the poor intestinal absorption and potential to cause ulcerations or strictures when bisphosphonates are taken orally.[13] Care should be taken to avoid hypocalcemia when using bisphosphonate therapy in bariatric surgery patients by ensuring adequate calcium and vitamin D intake.

As noted in the micronutrient deficiency section, inadequate levels of vitamin D are present in 80% to 90% of obese subjects prior to surgery.[31–33,49] This may be due to decreased rates of cutaneous vitamin D synthesis, decreased exposure to sunlight, and decreased intestinal absorption of ingested vitamin D in obese individuals.[50] Intestinal absorption of vitamin D is further decreased in malabsorptive procedures.[39] Vitamin D levels are commonly inadequate following bariatric surgery, despite supplementation.[38,39] Patients should be treated aggressively with vitamin D supplementation to ensure adequate stores for proper bone metabolism.[13]

Secondary hyperparathyroidism has been noted in 30% to 50% of RYGB patients.[39,51,52] This may be due to insufficient levels of vitamin D as well as a negative calcium balance. In one study, a single large dose of cholecalciferol was unable to reduce parathyroid hormone concentrations despite normalization of 25-hydroxyvitamin D levels, suggesting that the observed secondary hyperparathyroidism may be due to other factors, such as low calcium intake.[53] The deleterious effects of vitamin D insufficiency, negative calcium balance, and secondary hyperparathyroidism can lead to severe osteomalacia.[54,55] In addition to adequate vitamin D supplementation, bariatric surgery patients should be given 1200 to 1500 mg of calcium supplementation daily.[13] Following BPD procedures, at least 2000 mg total calcium should be given daily.[13] Because of decreased gastric acid production, calcium citrate is preferred due to superior bioavailability compared with other types of calcium supplements.[56,57]

NEPHROLITHIASIS

RYGB patients experience changes in oxalate metabolism that result in increased urinary oxalate excretion.[58,59] Daily urinary oxalate excretion increases by 25% to 32% after surgery.[58,59] Simultaneously, urinary citrate levels decrease by 32% following RYGB.[59] Both of these factors increase the risk of urinary oxalate stone formation. A relatively high prevalence of oxalate nephrolithiasis (4%) has been noted in one series of RYGB patients, with an average time to stone formation of 2.9 years following surgery.[60] One retrospective study showed a relative risk of 1.6 for the incidence of nephrolithiasis in RYGB patients compared with obese controls.[61] Several reported case series describe the development of oxalate nephropathy in post-RYGB patients, resulting in renal failure and death in some cases.[62,63]

Currently, data from large trials confirming these findings are not available. Nevertheless, it would be prudent to take precautionary measures that reduce the risk of oxalate nephrolithiasis in RYGB patients. These patients should already be taking calcium supplementation for bone health, which will also reduce urinary oxalate excretion. Periodic 24-hour urinary oxalate and citrate levels should be measured. If citrate is low, supplementation with potassium citrate is recommended to reduce stone formation.[13] Patients should also follow a low-oxalate diet and avoid dehydration.

Presently, high urinary oxalate has only been reported in RYGB patients. One study suggests that LAGB surgery does not cause an increase in oxalate excretion.[64] Further studies are necessary before recommendations can be made regarding the need for prevention of nephrolithiasis in the bariatric surgery population.

DIABETES

Dramatic improvement in T2DM has been widely observed following bariatric surgery.[7] In this section we review the degree and time course of diabetes resolution after bariatric procedures. We also discuss factors influencing improvement of T2DM and possible mechanisms through which this occurs. Finally, we provide recommendations for perioperative management of diabetes medications and continued long-term care.

The profound effect of bariatric surgery on T2DM was originally described more than 20 years ago by Pories and colleagues, who found that 98% of patients with T2DM or impaired fasting glucose became euglycemic within 10 days of RYGB surgery.[65] Since then, multiple retrospective[66–71] and prospective cohort[72–74] studies, as well as some prospective control trials,[75–77] have reported high rates of resolution and improvement of T2DM after bariatric surgery. In the majority of studies, resolution is defined as some combination of returning to a fasting blood glucose <110 mg/dL or a glycated hemoglobin A_{1c} (HbA_{1c}) <7% without diabetic medications; improvement is defined as approaching these values or reaching these targets with less medication. The short follow-up time (usually <2 years) in most of these studies makes it difficult to establish the durability of these effects; however, there is evidence that improvement in diabetes is sustained for 5–20 years.[66,68–70,72] One retrospective analysis of 330 patients with T2DM or impaired fasting glucose who underwent RYGB reported normalization of glucose, insulin, and HbA_{1c} in 91% of patients for as long as 14 years.[66] High resolution rates are reported for up to 20 years in BPD patients.[70]

In contrast, these relatively high T2DM resolution rates may diminish over time. The Swedish Obese Subjects Study evaluated incidence of T2DM in surgically treated obese subjects compared with contemporaneously matched, conventionally treated controls at 2 and 10 years.[75] More than 4000 patients treated with either RYGB (13%), VBG (68%), GB (19%), or nonsurgical management were available for follow-up at 2 years, and >1700 were available at 10 years. The initial T2DM resolution rate among surgically treated patients was 72% at 2 years, but this decreased to 36% at 10 years. Cautious interpretation of these results is required due to the high attrition rate at 10 years. A decline in T2DM resolution rates from 82% at 2 years and 62% after 2 years was also observed in a large meta-analysis of studies specifically reporting on diabetic patients who underwent bariatric surgery.[8] Patients that undergo bariatric surgery should be aware that improvements in T2DM may be temporary and long-term follow-up is required.

The type of bariatric surgery is important in predicting postoperative T2DM resolution. Overall, malabsorptive surgeries have higher T2DM resolution rates than purely restrictive procedures. Initial T2DM resolution rates of 72%–98% for RYGB[65,72,73,75] and >90% for BPD[69,78,79] have been reported. T2DM resolution after gastric banding procedures is more variable, ranging from 40% to 80%.[71,76,77,80] A recent meta-analysis[8] revealed a complete resolution of T2DM in 57% of patients who underwent LAGB, 80% who underwent RYGB, and 95% who underwent BPD. Interestingly, although primarily a restrictive procedure, T2DM resolution rates after SG are similar to those for RYGB, mainly in the 80% range.[71,74]

The type of surgery also affects the timing of T2DM amelioration. Resolution of T2DM occurs within days to weeks of RYGB and BPD.[65,66,69,70,72] After RYGB, 30% of patients were able to discontinue diabetic medications by hospital discharge.[72] In contrast, studies on gastric restrictive procedures report T2DM improves after months, when substantial weight loss is achieved.[76,77] In a retrospective analysis, the time to T2DM resolution was 12.6 months for LAGB patients versus 3.2 months for RYGB and 3.3 months for SG.[71] By these results, patients undergoing malabsorptive procedures (RYGB or BPD) and perhaps SG are more likely to show early improvement in T2DM and should be closely monitored in the postoperative period for T2DM resolution so that medications can properly be adjusted. It should be noted that T2DM remission can be scattered over a wide time range, occurring before significant weight loss in some cases or concurrent with weight loss in others, regardless of procedure. Prospective studies specifically evaluating the time course of T2DM improvement are warranted.

Patient-specific prognostic factors for T2DM resolution have been identified (see Table 3). Several studies demonstrate that younger age, smaller preoperative waist circumference, shorter duration of T2DM, and lower preoperative therapeutic requirements are associated with increased rates of recovery from T2DM following RYGB surgery.[66,67,72,73] The magnitude of postoperative weight loss also correlates with recovery from T2DM (Table 3).[72]

Table 3. *Preoperative Factors Associated with Improvement in Diabetes After Bariatric Surgery*.*

Factor (Prior to Bariatric Surgery)	Improvement	Complete Resolution
Age	<48 years	<48 years
Insulin use	No	No
Duration of T2DM	<10 years	<4 years
HbA_{1c}	<8.8%	<8.0%
Waist circumference	<53 in.	<48 in.

Abbreviations: HbA_{1c}, glycated hemoglobin A_{1c}; T2DM, type 2 diabetes mellitus.

*Table based on reported data.[66,72,73]

Bariatric surgery improves T2DM through multiple mechanisms that include dietary, metabolic, and hormonal factors. Bariatric surgical patients consume fewer calories and simple carbohydrates and experience weight loss, all of which improve T2DM.[10] These factors, although beneficial, cannot fully explain the superiority of RYGB and BPD versus equivalent weight loss from dieting or restrictive operations, where caloric intake is also reduced. Postoperative changes in hormonal signal pathways that mediate glucose-energy metabolism also play an important and possibly dominant role in the resolution of T2DM following bariatric surgery.[7,81]

Improvements in insulin resistance, as well as restoration of the first-phase insulin response, have been shown within a month after BPD,[82] BPD with RYGB,[83] and RYGB alone.[84,85] These changes were not seen in patients who underwent gastric restrictive procedures, indicating that the rearrangement of the intestine is partly responsible for these alterations in beta-cell function and insulin sensitivity.[84]

The role of gut hormones as mediators of glucose metabolism after bariatric operations is actively being researched and is discussed in detail by Korner and Inabnet in this issue of the *Mount Sinai Journal of Medicine*. The role of ghrelin, an appetite-stimulating and prodiabetic hormone produced by the stomach and duodenum, is controversial.[86] Some studies have reported a decrease after RYGB and SG, though others have shown no change.[87,88] Meal-stimulated peptide YY increases satiety and delays gastric emptying, indirectly improving glycemic control. Peptide YY levels increase after BPD and RYGB, and interestingly have also been shown to increase in some studies after SG as well.[86]

The incretin hormones, which include glucagon-like peptide 1 (GLP-1) and glucose-dependent insulinotropic polypeptide, enhance insulin secretion, decrease glucagon secretion, inhibit gastric emptying, and exert trophic effects on beta cells in response to meals, are increased by SG, RYGB, and

BPD.[82,86,89] All of these effects can improve glucose metabolism in T2DM.

Targeted glycemic control before and after surgery is recommended and is associated with fewer infectious complications across a variety of surgical procedures.[13,90] Preoperatively, glycemic control should target a HbA_{1c} <7%, fasting blood glucose levels of 80–130 mg/dL, and postprandial blood glucose levels <160 mg/dL.[91] Patients awaiting bariatric surgery can attempt to achieve these goals via a combination of lifestyle changes, oral hypoglycemics, and insulin.

Diabetes can drastically and quickly improve in the postoperative period, and diabetes medications should be administered accordingly. Medication dose reduction may be necessary starting several days prior to surgery, coordinated with the change to liquid diets, which are commonly started before surgery. In our institution, oral hypoglycemic agents and mealtime insulin are held and basal insulin is reduced by half once the liquid diet is initiated up until surgery. Immediately after surgery, basal insulin is held and blood glucose is monitored every 6 hours. Rapid-acting insulin correction is given as needed for elevated blood glucose (>180 mg/dL). If blood glucose values are persistently elevated, weight-based basal insulin (and if eating, prandial insulin) is started. Current recommended glycemic targets include fasting blood glucose levels of 80–110 mg/dL and postprandial values <180 mg/dL in non–intensive care patients.[9] Blood glucose should be maintained at 80–110 mg/dL for postoperative patients in the intensive care unit with an intravenous insulin infusion.[9]

Diabetes can drastically and quickly improve in the postoperative period, and diabetes medications should be administered accordingly.

After hospital discharge, the decision to continue insulin, change to oral hypoglycemic agents, or stop all medications is made on an individual basis and is usually based on the preoperative diabetes regimen, as well as the degree of insulin requirements postoperatively. Younger patients, patients with less-severe T2DM, and patients with T2DM of relatively short duration may not require diabetic medications postoperatively. The type of procedure should also guide discharge medication management. Patients who undergo BPD, RYGB, and SG should have close monitoring for medication dose adjustments; medication requirements may change within the first postoperative month. All patients should monitor blood glucose closely and contact their physician if their blood glucose is persistently <70 or >200 mg/dL.

After hospital discharge, the decision to continue insulin, change to oral hypoglycemic agents, or stop all medications is made on an individual basis and is usually based on the preoperative diabetes regimen and the type of procedure, as well as the degree of insulin requirements postoperatively. Younger patients, patients with less-severe type 2 diabetes mellitus, and patients with diabetes of relatively short duration may not require diabetic medications postoperatively.

The choice of oral diabetes medication(s) poses additional challenges. Sulfonylureas and meglitinides can cause hypoglycemia, especially if patients are not eating regularly. Gastrointestinal discomfort after bariatric surgery may be worsened by metformin, α-glucosidase inhibitors, and GLP-1 analogs. Thiazolidinediones (TZD) may be safe in the short term in absence of any other contraindications. However, in light of recent reports of increased fracture risk with these agents,[92] caution is advised with long-term use in this population that is already at high risk for alterations in bone metabolism.

Over the first postoperative year, patients meet with the bariatric team at months 1, 3, 6, and 12 in our institution. HbA_{1c} measurements at these visits help to screen for T2DM recurrence in those who experienced recovery as well as to guide further medication management in those whose diabetes has not resolved. After 12 months, persistently diabetic patients should be managed by a primary physician and/or endocrinologist according to best standard of care.[91] In those whose T2DM has remitted, primary care physicians should periodically screen for recurrence. Patients should be counseled regarding symptoms of hyperglycemia and understand that their diabetes may return, especially if they regain weight after the initial weight loss.

POLYCYSTIC OVARIAN SYNDROME AND FEMALE FERTILITY

PCOS is defined by the presence of chronic anovulation and androgen excess and affects approximately 6% of premenopausal women in the United States.[93] Women present with varying degrees of virilization and infertility. PCOS is highly associated with insulin resistance and the metabolic syndrome.[94] Women with PCOS gain weight easily, and weight gain tends to worsen symptoms of PCOS. The underlying pathophysiology of this disorder is subject to debate. Many theorize that insulin resistance, possibly at the level of the ovary, as well as disruption of the hypothalamic-pituitary-ovarian axis may be responsible.[95] Treatments for PCOS include the use of insulin-sensitizing agents accompanied by weight loss and healthy lifestyle choices. These may restore ovulatory cycles and improve fertility.[96]

Bariatric surgery is associated with improved insulin sensitivity and weight loss and effectively resolves PCOS in adults[97,98] and adolescents.[99] In one retrospective study of 24 obese women with PCOS, 75% showed moderate to complete resolution of hirsutism over a period of 6–49 months following RYGB.[98] Normal menstruation returned in 100% of patients within a mean of 3.4 ± 2.1 months and a range of 1–10 months. Of the 5 patients with preoperative infertility, all were able to conceive postoperatively.

Similar results were reported in a prospective nonrandomized study of 12 morbidly obese women with PCOS, all of whom regained regular menstrual cycles after BPD or RYGB.[97] In this study, insulin sensitivity, as measured by homeostasis model assessment, improved and there were significant decreases in biochemical and clinical signs of hyperandrogenism in all but 1 patient after a mean of 12 ± 5 months.

Women who undergo bariatric surgery should understand that menstrual irregularities will likely normalize. One survey-based study showed recovery of menstrual function in 70 of 98 anovulatory women after bariatric surgery.[100] Menstruation returned on average 1.6 months postoperatively. There was greater absolute weight loss in the women who regained menstrual function. Although this study did not comment on the effect of bariatric surgery on fertility rates, other studies have reported increased spontaneous conception.[98,101,102] Of note, the efficacy of oral contraceptive agents may decrease after malabsorptive procedures due to decreased bioavailability.[103]

Whether there is an increased risk of pregnancy complications following bariatric surgery is controversial. Based on expert opinion, the American College of Obstetrics and Gynecology advises women to delay pregnancy for 12–18 months after bariatric surgery.[16] There are early reports of maternal complications, as well as micronutrient deficiencies leading to fetal defects.[104] However, bariatric surgery reduces the risks for gestational diabetes, preeclampsia, macrosomia, and caesarean delivery in the long term.[105] In addition, decreased childhood obesity after maternal bariatric surgery has been reported.[106] A cross-sectional study evaluating children born to mothers before and after BPD showed a 3-fold reduction in severe obesity in children born after weight-loss surgery.[106] These children also demonstrated more favorable lipid profiles and lower markers of insulin resistance.

The decision to become pregnant should be made on an individual basis, and women should be counseled about the importance of nutritional supplementation and dietary adherence. Pregnant women with LAGB need to be closely monitored during pregnancy, as adjustments to the band may be required.[107]

MALE SEX HORMONES AND FERTILITY

Obesity is associated with decreased sexual quality of life, erectile dysfunction, alterations in sperm parameters, and infertility in males.[108–110] This may be partly due to changes in sex hormones that are observed in obese men, including decreased total and free testosterone, decreased sex hormone binding globulin, and increased estrogen. The prevalence of hypoandrogenism in obese males has been reported as 20%–64%.[108] In many studies, nonsurgical weight loss improves both serum testosterone levels and clinical measures of men's sexual health.[111–115] Similar effects on gonadal function are noted following weight loss after bariatric surgery in men.[91,116–120]

In one study, total testosterone levels returned to the normal range from subnormal preoperative values at 1 year after BPD.[117] Serum estradiol also decreased to normal male levels.[117] Comparable results have been demonstrated after VBG[118] and RYGB.[116] A cohort study evaluating sex hormones 2 years after RYGB showed increased total and free testosterone, increased SHBG, and decreased estradiol.[116] Moreover, sexual quality of life, assessed by a questionnaire including information regarding libido and performance, was significantly increased. Our own unpublished data support these findings.[119]

Improvement in erectile dysfunction has been shown to correlate with the degree of weight loss after RYGB.[120] Patients should be informed regarding these changes in libido and sexual function that they may experience postoperatively.

Little is known about the effect bariatric surgery has on male fertility. Given the improvement in the sex-hormone profile, generally one would expect increased fertility after bariatric surgery. However, a few cases of infertility following RYGB in a previously fertile male have been reported.[121] This may be the result of malnutrition. Overall, improvements in testosterone levels mirror the multiple metabolic improvements that occur following bariatric surgery and are likely beneficial to male sexual health and reproductive function.[120]

SUMMARY

Bariatric surgical procedures induce substantial weight loss and improve obesity-related conditions, providing great benefits to obese patients. Consideration for the nutritional, metabolic, and hormonal changes that occur in the postoperative period can lead to favorable patient outcomes. Our recommendations for postoperative patient care are summarized below.

- Postoperative diet should be advanced carefully (Table 1).
- Patients should be screened for preoperative and postoperative malnutrition, including protein-calorie malnutrition in malabsorptive procedures and micronutrient deficiencies in all bariatric surgeries.
- For improved bone health, calcium and vitamin D supplementation should be administered in all patients.
- Patients should be screened for oxalate excretion and risk of nephrolithiasis.
- High rates of T2DM resolution are noted within 1 month for RYGB, BPD, and SG procedures and with weight loss following restrictive procedures. Diabetic patients should be closely monitored and regimen adjusted as necessary.
- PCOS and female infertility improve postoperatively.
- In men, testosterone levels increase and hypogonadism associated with obesity often resolves.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999–2008. *JAMA* 2010; 303: 235–241.
2. Adams TD, Gress RE, Smith SC, et al. Long-term mortality after gastric bypass surgery. *N Engl J Med* 2007; 357: 753–761.
3. Sjöström L, Narbro K, Sjöström CD, et al. Effects of bariatric surgery on mortality in Swedish obese subjects. *N Engl J Med* 2007; 357: 741–752.
4. Sjöström L. Bariatric surgery and reduction in morbidity and mortality: experiences from the SOS study. *Int J Obes (Lond)* 2008; 32(suppl 7): S93–S97.
5. Gastrointestinal surgery for severe obesity: National Institutes of Health Consensus Development Conference Statement. *Am J Clin Nutr* 1992; 55(suppl 2): 615S–619S.
6. Santry HP, Gillen DL, Lauderdale DS. Trends in bariatric surgical procedures. *JAMA* 2005; 294: 1909–1917.
7. Herron DM, Tong W. Role of surgery in management of type 2 diabetes mellitus. *Mt Sinai J Med* 2009; 76: 281–293.
8. Buchwald H, Estok R, Fahrbach K, et al. Weight and type 2 diabetes after bariatric surgery: systematic review and meta-analysis. *Am J Med* 2009; 122: 248–256.e5.
9. Mechanick JI, Kushner RF, Sugerman HJ, et al. American Association of Clinical Endocrinologists, The Obesity Society, and American Society for Metabolic & Bariatric Surgery medical guidelines for clinical practice for the perioperative nutritional, metabolic, and nonsurgical support of the bariatric surgery patient. *Obesity (Silver Spring)* 2009; 17(suppl 1): S1–S70, v.
10. Stocker DJ. Management of the bariatric surgery patient. *Endocrinol Metab Clin North Am* 2003; 32: 437–457.
11. Ribeiro AG, Costa MJ, Faintuch J, Dias MC. A higher meal frequency may be associated with diminished weight loss after bariatric surgery. *Clinics (Sao Paulo)* 2009; 64: 1053–1058.
12. Apovian CM, Cummings S, Anderson W, et al. Best practice updates for multidisciplinary care in weight loss surgery. *Obesity (Silver Spring)* 2009; 17: 871–879.
13. Mechanick JI, Kushner RF, Sugerman HJ, et al. American Association of Clinical Endocrinologists, The Obesity Society, and American Society for Metabolic & Bariatric Surgery Medical Guidelines for Clinical Practice for the perioperative nutritional, metabolic, and nonsurgical support of the bariatric surgery patient [published correction appears in *Surg Obes Relat Dis* 2010; 6: 112]. *Surg Obes Relat Dis* 2008; 4(5 suppl): S109–S184.
14. Mallory GN, Macgregor AM, Rand CS. The influence of dumping on weight loss after gastric restrictive surgery for morbid obesity. *Obes Surg* 1996; 6: 474–478.
15. Tack J, Arts J, Caenepeel P, et al. Pathophysiology, diagnosis and management of postoperative dumping syndrome. *Nat Rev Gastroenterol Hepatol* 2009; 6: 583–590.
16. Bloomberg RD, Fleishman A, Nalle JE, et al. Nutritional deficiencies following bariatric surgery: what have we learned? *Obes Surg* 2005; 15: 145–154.

17. Malinowski SS. Nutritional and metabolic complications of bariatric surgery. *Am J Med Sci* 2006; 331: 219–225.
18. Faintuch J, Matsuda M, Cruz ME, et al. Severe protein-calorie malnutrition after bariatric procedures. *Obes Surg* 2004; 14: 175–181.
19. Brolin RE, LaMarca LB, Kenler HA, Cody RP. Malabsorptive gastric bypass in patients with superobesity. *J Gastrointest Surg* 2002; 6: 195–205.
20. Skroubis G, Sakellaropoulos G, Pouggouras K, et al. Comparison of nutritional deficiencies after Roux-en-Y gastric bypass and after biliopancreatic diversion with Roux-en-Y gastric bypass. *Obes Surg* 2002; 12: 551–558.
21. Marceau S, Biron S, Lagacé M, et al. Biliopancreatic diversion. with distal gastrectomy, 250 cm and 50 cm limbs: long-term results. *Obes Surg* 1995; 5: 302–307.
22. Marinari GM, Murelli F, Camerini G, et al. A 15-year evaluation of biliopancreatic diversion according to the Bariatric Analysis Reporting Outcome System (BAROS). *Obes Surg* 2004; 14: 325–328.
23. Bock MA. Roux-en-Y gastric bypass: the dietitian's and patient's perspectives. *Nutr Clin Pract* 2003; 18: 141–144.
24. Alvarez-Leite JI. Nutrient deficiencies secondary to bariatric surgery. *Curr Opin Clin Nutr Metab Care* 2004; 7: 569–575.
25. Fujioka K. Follow-up of nutritional and metabolic problems after bariatric surgery. *Diabetes Care* 2005; 28: 481–484.
26. Ernst B. Thurnheer M, Schmid SM, Schultes B. Evidence for the necessity to systematically assess micronutrient status prior to bariatric surgery. *Obes Surg* 2009; 19: 66–73.
27. Kaidar-Person O, Person B, Szomstein S, Rosenthal RJ. Nutritional deficiencies in morbidly obese patients: a new form of malnutrition? Part A: vitamins. *Obes Surg* 2008; 18: 870–876.
28. Kaidar-Person O, Person B, Szomstein S, Rosenthal RJ. Nutritional deficiencies in morbidly obese patients: a new form of malnutrition? Part B: minerals. *Obes Surg* 2008; 18: 1028–1034.
29. Agte V, Tarwadi K, Mengale S, et al. Vitamin profile of cooked foods: how healthy is the practice of ready-to-eat foods? *Int J Food Sci Nutr* 2002; 53: 197–208.
30. Markison S. The role of taste in the recovery from specific nutrient deficiencies in rats. *Nutr Neurosci* 2001; 4: 1–14.
31. Abbasi AA, Amin M, Smiertka JK, et al. Abnormalities of vitamin D and calcium metabolism after surgical treatment of morbid obesity: a study of 136 patients. *Endocr Pract* 2007; 13: 131–136.
32. Bischof MG, Heinze G, Vierhapper H. Vitamin D status and its relation to age and body mass index. *Horm Res* 2006; 66: 211–215.
33. Carlin AM, Rao DS, Meslemani AM, et al. Prevalence of vitamin D depletion among morbidly obese patients seeking gastric bypass surgery. *Surg Obes Relat Dis* 2006; 2: 98–104.
34. Coupaye M, Puchaux K, Bogard C, et al. Nutritional consequences of adjustable gastric banding and gastric bypass: a 1-year prospective study. *Obes Surg* 2009; 19: 56–65.
35. De Luis DA, Pacheco D, Izaola O, et al. Clinical results and nutritional consequences of biliopancreatic diversion: three years of follow-up. *Ann Nutr Metab* 2008; 53: 234–239.
36. Marceau P, Hould FS, Lebel S, et al. Malabsorptive obesity surgery. *Surg Clin North Am* 2001; 81: 1113–1127.
37. Rhode BM, Tamin H, Gilfix BM, et al. Treatment of vitamin B12 deficiency after gastric surgery for severe obesity. *Obes Surg* 1995; 5: 154–158.
38. Clements RH, Katasani VG, Palepu R, et al. Incidence of vitamin deficiency after laparoscopic Roux-en-Y gastric bypass in a university hospital setting. *Am Surg* 2006; 72: 1196–1204.
39. Johnson JM, Maher JW, DeMaria EJ, et al. The long-term effects of gastric bypass on vitamin D metabolism. *Ann Surg* 2006; 243: 701–705.
40. Coates PS, Fernstrom JD, Fernstrom MH, et al. Gastric bypass surgery for morbid obesity leads to an increase in bone turnover and a decrease in bone mass. *J Clin Endocrinol Metab* 2004; 89: 1061–1065.
41. DiGiorgi M, Daud A, Inabnet WB, et al. Markers of bone and calcium metabolism following gastric bypass and laparoscopic adjustable gastric banding. *Obes Surg* 2008; 18: 1144–1148.
42. Goldner WS, O'Dorisio TM, Dillon JS, Mason EE. Severe metabolic bone disease as a long-term complication of obesity surgery. *Obes Surg* 2002; 12: 685–692.
43. Johnson JM, Maher JW, Samuel I, et al. Effects of gastric bypass procedures on bone mineral density, calcium, parathyroid hormone, and vitamin D. *J Gastrointest Surg* 2005; 9: 1106–1111.
44. Ott MT, Fanti P, Malluche HH, et al. Biochemical evidence of metabolic bone disease in women following Roux-Y gastric bypass for morbid obesity. *Obes Surg* 1992; 2: 341–348.
45. Fleischer J, Stein EM, Bessler M, et al. The decline in hip bone density after gastric bypass surgery is associated with extent of weight loss. *J Clin Endocrinol Metab* 2008; 93: 3735–3740.
46. Dixon JB, Strauss BJ, Laurie C, O'Brien PE. Changes in body composition with weight loss: obese subjects randomized to surgical and medical programs. *Obesity (Silver Spring)* 2007; 15: 1187–1198.
47. Strauss BJ, Marks SJ, Growcott JP, et al. Body composition changes following laparoscopic gastric banding for morbid obesity. *Acta Diabetol* 2003; 40(suppl 1): S266–S269.
48. El-Kadre LJ, Rocha PR, de Almeida Tinoco AC, Tinoco RC. Calcium metabolism in pre-and postmenopausal morbidly obese women at baseline and after laparoscopic Roux-en-Y gastric bypass. *Obes Surg* 2004; 14: 1062–1066.
49. Hamoui N, Anthone G, Crookes PF. Calcium metabolism in the morbidly obese. *Obes Surg* 2004; 14: 9–12.
50. Wortsman J, Matsuoka LY, Chen TC, et al. Decreased bioavailability of vitamin D in obesity. *Am J Clin Nutr* 2000; 72: 690–693.
51. Jin J, Robinson AV, Hallowell PT, et al. Increases in parathyroid hormone (PTH) after gastric bypass surgery appear to be of a secondary nature. *Surgery* 2007; 142: 914–920.
52. Youssef Y, Richards WO, Sekhar N, et al. Risk of secondary hyperparathyroidism after laparoscopic gastric bypass surgery in obese women. *Surg Endosc* 2007; 21: 1393–1396.

53. Einarsdóttir K, Preen DB, Clay TD, et al. Effect of a single 'megadose' intramuscular vitamin D (600,000 IU) injection on vitamin D concentrations and bone mineral density following biliopancreatic diversion surgery. *Obes Surg* 2010; 20: 732–737.
54. Collazo-Clavell ML, Jimenez A, Hodgson SF, Sarr MG. Osteomalacia after Roux-en-Y gastric bypass. *Endocr Pract* 2004; 10: 195–198.
55. De Prisco C, Levine SN. Metabolic bone disease after gastric bypass surgery for obesity. *Am J Med Sci* 2005; 329: 57–61.
56. Miller AD, Smith KM. Medication and nutrient administration considerations after bariatric surgery. *Am J Health Syst Pharm* 2006; 63: 1852–1857.
57. Recker RR. Calcium absorption and achlorhydria. *N Engl J Med* 1985; 313: 70–73.
58. Duffey BG, Pedro RN, Makhlouf A, et al. Roux-en-Y gastric bypass is associated with early increased risk factors for development of calcium oxalate nephrolithiasis. *J Am Coll Surg* 2008; 206: 1145–1153.
59. Park AM, Storm DW, Fulmer BR, et al. A prospective study of risk factors for nephrolithiasis after Roux-en-Y gastric bypass surgery. *J Urol* 2009; 182: 2334–2339.
60. Sinha MK, Collazo-Clavell ML, Rule A, et al. Hyperoxaluric nephrolithiasis is a complication of Roux-en-Y gastric bypass surgery. *Kidney Int* 2007; 72: 100–107.
61. Matlaga BR, Shore AD, Magnuson T, et al. Effect of gastric bypass surgery on kidney stone disease. *J Urol* 2009; 181: 2573–2577.
62. Nasr SH, D'Agati VD, Said SM, et al. Oxalate nephropathy complicating Roux-en-Y Gastric Bypass: an underrecognized cause of irreversible renal failure. *Clin J Am Soc Nephrol* 2008; 3: 1676–1683.
63. Nelson WK, Houghton SG, Milliner DS, et al. Enteric hyperoxaluria, nephrolithiasis, and oxalate nephropathy: potentially serious and unappreciated complications of Roux-en-Y gastric bypass. *Surg Obes Relat Dis* 2005; 1: 481–485.
64. Penniston KL, Kaplon DM, Gould JC, Nakada SY. Gastric band placement for obesity is not associated with increased urinary risk of urolithiasis compared to bypass. *J Urol* 2009; 182: 2340–2346.
65. Pories WJ. Caro JF. Flickinger EG, et al. The control of diabetes mellitus (NIDDM) in the morbidly obese with the Greenville Gastric Bypass. *Ann Surg* 1987; 206: 316–323.
66. Pories WJ, Swanson MS, MacDonald KG, et al. Who would have thought it? An operation proves to be the most effective therapy for adult-onset diabetes mellitus. *Ann Surg* 1995; 222: 339–352.
67. Sugerman HJ, Wolfe LG, Sica DA, Clore JN. Diabetes and hypertension in severe obesity and effects of gastric bypass-induced weight loss. *Ann Surg* 2003; 237: 751–758.
68. MacDonald KG Jr., Long SD, Swanson MS, et al. The gastric bypass operation reduces the progression and mortality of non-insulin-dependent diabetes mellitus. *J Gastrointest Surg* 1997; 1: 213–220.
69. Scopinaro N, Marinari GM, Camerini GB, et al. Specific effects of biliopancreatic diversion on the major components of metabolic syndrome: a long-term follow-up study. *Diabetes Care* 2005; 28: 2406–2411.
70. Scopinaro N, Papadia F, Camerini G, et al. A comparison of a personal series of biliopancreatic diversion and literature data on gastric bypass help to explain the mechanisms of resolution of type 2 diabetes by the two operations. *Obes Surg* 2008; 18: 1035–1038.
71. Abbatini F, Rizzello M, Casella G, et al. Long-term effects of laparoscopic sleeve gastrectomy, gastric bypass, and adjustable gastric banding on type 2 diabetes. *Surg Endosc* 2010; 24: 1005–1010.
72. Schauer PR, Burguera B, Ikramuddin S, et al. Effect of laparoscopic Roux-en Y gastric bypass on type 2 diabetes mellitus. *Ann Surg* 2003; 238: 467–485.
73. Torquati A, Lutfi R, Abumrad N, Richards WO. Is Roux-en-Y gastric bypass surgery the most effective treatment for type 2 diabetes mellitus in morbidly obese patients? *J Gastrointest Surg* 2005; 9: 1112–1118.
74. Vidal J, Ibarzabal A, Romero F, et al. Type 2 diabetes mellitus and the metabolic syndrome following sleeve gastrectomy in severely obese subjects. *Obes Surg* 2008; 18: 1077–1082.
75. Sjöström L, Lindroos AK, Peltonen M, et al. Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. *N Engl J Med* 2004; 351: 2683–2693.
76. Dixon JB, O'Brien PE, Playfair J, et al. Adjustable gastric banding and conventional therapy for type 2 diabetes: a randomized controlled trial. *JAMA* 2008; 299: 316–323.
77. O'Brien PE, Dixon JB, Laurie C, et al. Treatment of mild to moderate obesity with laparoscopic adjustable gastric banding or an intensive medical program: a randomized trial. *Ann Intern Med* 2006; 144: 625–633.
78. Marceau P, Biron S, Hould FS, et al. Duodenal switch: long-term results. *Obes Surg* 2007; 17: 1421–1430.
79. Marinari GM, Papadia FS, Briatore L, et al. Type 2 diabetes and weight loss following biliopancreatic diversion for obesity. *Obes Surg* 2006; 16: 1440–1444.
80. Tice JA, Karliner L, Walsh J, et al. Gastric banding or bypass? A systematic review comparing the two most popular bariatric procedures. *Am J Med* 2008; 121: 885–893.
81. Lee WJ, Wang W, Lee YC, et al. Effect of laparoscopic mini-gastric bypass for type 2 diabetes mellitus: comparison of BMI >35 and <35 kg/m^2. *J Gastrointest Surg* 2008; 12: 945–952.
82. Salinari S, Bertuzzi A, Asnaghi S, et al. First-phase insulin secretion restoration and differential response to glucose load depending on the route of administration in type 2 diabetic subjects after bariatric surgery. *Diabetes Care* 2009; 32: 375–380.
83. Polyzogopoulou EV, Kalfarentzos F, Vagenakis AG, Alexandrides TK. Restoration of euglycemia and normal acute insulin response to glucose in obese subjects with type 2 diabetes following bariatric surgery. *Diabetes* 2003; 52: 1098–1103.
84. Kashyap SR, Daud S, Kelly KR, et al. Acute effects of gastric bypass versus gastric restrictive surgery on beta-cell function and insulinotropic hormones in severely obese patients with type 2 diabetes. *Int J Obes (Lond)* 2010; 34: 462–471.
85. Wickremesekera K, Miller G, Naotunne TD, et al. Loss of insulin resistance after Roux-en-Y gastric bypass surgery: a time course study. *Obes Surg* 2005; 15: 474–481.
86. Thaler JP, Cummings DE. Minireview: Hormonal and metabolic mechanisms of diabetes remission after

gastrointestinal surgery. *Endocrinology* 2009; 150: 2518–2525.

87. Karamanakos SN, Vagenas K, Kalfarentzos F, Alexandrides TK. Weight loss, appetite suppression, and changes in fasting and postprandial ghrelin and peptide-YY levels after Roux-en-Y gastric bypass and sleeve gastrectomy: a prospective, double blind study. *Ann Surg* 2008; 247: 401–407.
88. Cummings DE, Weigle DS, Frayo RS, et al. Plasma ghrelin levels after diet-induced weight loss or gastric bypass surgery. *N Engl J Med* 2002; 346: 1623–1630.
89. Peterli R, Wölnerhanssen B, Peters T, et al. Improvement in glucose metabolism after bariatric surgery: comparison of laparoscopic Roux-en-Y gastric bypass and laparoscopic sleeve gastrectomy: a prospective randomized trial. *Ann Surg* 2009; 250: 234–241.
90. Dronge AS, Perkal MF, Kancir S, et al. Long-term glycemic control and postoperative infectious complications. *Arch Surg* 2006; 141: 375–380.
91. Reis LO, Favaro WJ, Barreiro GC, et al. Erectile dysfunction and hormonal imbalance in morbidly obese male is reversed after gastric bypass surgery: a prospective randomized controlled trial. *Int J Androl* 2009;doi:10.1111/j.1365-2605.2010.01097.x.
92. Habib ZA, Havstad SL, Wells K, et al. Thiazolidinedione use and the longitudinal risk of fractures in patients with type 2 diabetes mellitus. *J Clin Endocrinol Metab* 2010; 95: 592–600.
93. Franks S, White DM. Prevalence of and etiological factors in polycystic ovarian syndrome. *Ann N Y Acad Sci* 1993; 687: 112–114.
94. Apridonidze T, Essah PA, Iuorno MJ, Nestler JE. Prevalence and characteristics of the metabolic syndrome in women with polycystic ovary syndrome. *J Clin Endocrinol Metab* 2005; 90: 1929–1935.
95. Kronenberg H, Williams RH. *Williams Textbook of Endocrinology*. 11th ed. Philadelphia, PA: Saunders/Elsevier; 2008.
96. Hoeger KM, Kochman L, Wixom N, et al. A randomized, 48-week, placebo-controlled trial of intensive lifestyle modification and/or metformin therapy in overweight women with polycystic ovary syndrome: a pilot study. *Fertil Steril* 2004; 82: 421–429.
97. Escobar-Morreale HF, Botella-Carretero JI, Alvarez-Blasco F, et al. The polycystic ovary syndrome associated with morbid obesity may resolve after weight loss induced by bariatric surgery. *J Clin Endocrinol Metab* 2005; 90: 6364–6369.
98. Eid GM, Cottam DR, Velcu LM, et al. Effective treatment of polycystic ovarian syndrome with Roux-en-Y gastric bypass. *Surg Obes Relat Dis* 2005; 1: 77–80.
99. Collins J, Mattar S, Qureshi F, et al. Initial outcomes of laparoscopic Roux-en-Y gastric bypass in morbidly obese adolescents. *Surg Obes Relat Dis* 2007; 3: 147–152.
100. Teitelman M, Grotegut CA, Williams NN, Lewis JD. The impact of bariatric surgery on menstrual patterns. *Obes Surg* 2006; 16: 1457–1463.
101. Bilenka B, Ben-Shlomo I, Cozacov C, et al. Fertility, miscarriage and pregnancy after vertical banded gastroplasty operation for morbid obesity. *Acta Obstet Gynecol Scand* 1995; 74: 42–44.
102. Deitel M, Stone E, Kassam HA, et al. Gynecologic-obstetric changes after loss of massive excess weight following bariatric surgery. *J Am Coll Nutr* 1988; 7: 147–153.
103. Merhi ZO. Impact of bariatric surgery on female reproduction. *Fertil Steril* 2009; 92: 1501–1508.
104. Haddow JE, Hill LE, Kloza EM, Thanhauser D. Neural tube defects after gastric bypass. *Lancet* 1986; 1: 1330.
105. Maggard MA, Yermilov I, Li Z, et al. Pregnancy and fertility following bariatric surgery: a systematic review. *JAMA* 2008; 300: 2286–2296.
106. Smith J, Cianflone K, Biron S, et al. Effects of maternal surgical weight loss in mothers on intergenerational transmission of obesity. *J Clin Endocrinol Metab* 2009; 94: 4275–4283.
107. Martin LF, Finigan KM, Nolan TE. Pregnancy after adjustable gastric banding. *Obstet Gynecol* 2000; 95: 927–930.
108. Diaz-Arjonilla M, Schwarcz M, Swerdloff RS, Wang C. Obesity, low testosterone levels and erectile dysfunction. *Int J Impot Res* 2009; 21: 89–98.
109. Hammoud AO, Wilde N, Gibson M, et al. Male obesity and alteration in sperm parameters. *Fertil Steril* 2008; 90: 2222–2225.
110. Sallmen M, Sandler DP, Hoppin JA, et al. Reduced fertility among overweight and obese men. *Epidemiology* 2006; 17: 520–523.
111. Leenen R, van der Kooy K, Seidell JC, et al. Visceral fat accumulation in relation to sex hormones in obese men and women undergoing weight loss therapy. *J Clin Endocrinol Metab* 1994; 78: 1515–1520.
112. Niskanen L, Laaksonen DE, Punnonen K, et al. Changes in sex hormone-binding globulin and testosterone during weight loss and weight maintenance in abdominally obese men with the metabolic syndrome. *Diabetes Obes Metab* 2004; 6: 208–215.
113. Kaukua J, Pekkarinen T, Sane T, Mustajoki P. Sex hormones and sexual function in obese men losing weight. *Obes Res* 2003; 11: 689–694.
114. Esposito K, Giugliano F, Di Palo C, et al. Effect of lifestyle changes on erectile dysfunction in obese men: a randomized controlled trial. *JAMA* 2004; 291: 2978–2984.
115. Esposito K, Ciotola M, Giugliano F, et al. Effects of intensive lifestyle changes on erectile dysfunction in men. *J Sex Med* 2009; 6: 243–250.
116. Hammoud A, Gibson M, Hunt SC, et al. Effect of Roux-en-Y gastric bypass surgery on the sex steroids and quality of life in obese men. *J Clin Endocrinol Metab* 2009; 94: 1329–1332.
117. Alagna S, Cossu ML, Gallo P, et al. Biliopancreatic diversion: long-term effects on gonadal function in severely obese men. *Surg Obes Relat Dis* 2006; 2: 82–86.
118. Bastounis EA, Karayiannakis AJ, Syrigos K, et al. Sex hormone changes in morbidly obese patients after vertical banded gastroplasty. *Eur Surg Res* 1998; 30: 43–47.
119. Omana JJ, Tamler R, Strohmayer E, et al. Sex hormone levels in men undergoing bariatric surgery. *J Am Coll Surg* 2009; 209(suppl 1): S22–S23.
120. Dallal RM, Chernoff A, O'Leary MP, et al. Sexual dysfunction is common in the morbidly obese male and improves after gastric bypass surgery. *J Am Coll Surg* 2008; 207: 859–864.

121. Di Frega AS, Dale B, Di Matteo L, Wilding M. Secondary male factor infertility after Roux-en-Y gastric bypass for morbid obesity: case report. *Hum Reprod* 2005; 20: 997–998.
122. Gehrer S, Kern B, Peters T, et al. Fewer nutrient deficiencies after laparoscopic sleeve gastrectomy (LSG) than after laparoscopic Roux-Y-gastric bypass (LRYGB): a prospective study. *Obes Surg* 2010; 20: 447–453.
123. Ledoux S, Msika S, Moussa F, et al. Comparison of nutritional consequences of conventional therapy of obesity, adjustable gastric banding, and gastric bypass. *Obes Surg* 2006; 16: 1041–1049.

Hormone Changes Affecting Energy Homeostasis after Metabolic Surgery

E. John Harvey, MD,[1] Kervin Arroyo, MD,[1] Judith Korner, MD, PhD,[2] and William B. Inabnet, MD, FACS, MD, FACS[1]

[1]Mount Sinai School of Medicine, New York, NY
[2]Columbia University School of Medicine, New York, NY

OUTLINE

Obesity is a Metabolic Disease
Defending a Weight Set Point: Mechanisms of Energy Homeostasis
  Central Nervous System Control of Feeding
  Peripheral Feedback Signals
The Hormonal Basis of Metabolic Surgery
  Ghrelin
  Glucagon-like Peptide-1
  Peptide Tyrosine Tyrosine
  Insulin
  Leptin
  Adiponectin
Tailored Metabolic Operations
Combining Metabolic Surgery and Pharmacotherapy to Optimize Therapy
Conclusion

ABSTRACT

Worldwide, there is an epidemic of obesity and its associated diseases. The past decade of research has brought about a paradigm shift in our understanding of both the mechanisms underlying energy homeostasis and the multiple factors contributing to the pathophysiology of obesity. Metabolic surgery is currently far more effective than diet and exercise or pharmacotherapy in achieving durable weight loss. Moreover, the remarkable results of surgery in achieving a rapid remission of type 2 diabetes mellitus has sparked tremendous excitement and research into the mechanisms through which metabolic surgery has its dramatic effect. As opposed to the traditional understanding of "restriction" and "malabsorption," current evidence suggests that metabolic surgery alters the expression of multiple hormones that affect both short-term and long-term signals of energy balance. We review the hormonal changes following the most common types of metabolic operations currently being performed. The profile of hormonal changes provides a guide to tailor the choice of operation for each individual patient toward achieving the desired metabolic result. In the future, individualized metabolic surgery alone or modulated by targeted pharmacological therapy may achieve the most reliable and effective results with the highest safety and lowest side effect profile. ***Mt Sinai J Med 77:446–465, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: adiponectin, bariatric, diabetes, ghrelin, GLP-1, insulin, leptin, metabolic, PYY, surgery.

Address Correspondence to:

William B. Inabnet, MD, FACS
Department of
Surgery–Division of Metabolic,
Endocrine, and Minimally
Invasive Surgery, Mount Sinai
Medical Center, New York, NY
Email: william.inabnet@
mountsinai.org

Obesity rates in the United States have been steadily rising since the 1970s, according to data from the Centers for Disease Control and Prevention.[1] In the 2007–2008 National Health and Nutrition Examination Survey, 33.8% of the US population was found to be obese (body mass index [BMI] >30 kg/m^2), with 5.7% having grade 3 obesity (BMI >40 kg/m^2). Worldwide, obesity rates have also been increasing, and the World Health Organization has predicted that by 2015 more than 700 million people will be obese.

OBESITY IS A METABOLIC DISEASE

The adverse health effects of obesity are well known.[2] Hypertension, hyperlipidemia, heart disease, type 2

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20203

© 2010 Mount Sinai School of Medicine

diabetes mellitus (T2DM), obstructive sleep apnea, thromboembolic disease, degenerative joint disease, gastroesophageal reflux disease, asthma, polycystic ovary disease, nonalcoholic fatty liver disease, as well as several cancers are all strongly associated with obesity. The rapid acceptance of the term "metabolic syndrome" to describe the association of truncal obesity with hypertension, hyperlipidemia, glucose intolerance, and insulin resistance highlights the importance of excess adiposity in causing metabolic imbalance.[3] Insulin resistance and chronic low-grade inflammation appear to be key elements of this abnormal physiologic state.

> *The rapid acceptance of the term "metabolic syndrome" to describe the association of truncal obesity with hypertension, hyperlipidemia, glucose intolerance, and insulin resistance highlights the importance of excess adiposity in causing metabolic imbalance.*

Multiple factors contribute to the development of obesity, including social factors such as the availability and affordability of calorie-dense foods (increased energy intake) and trends toward a sedentary lifestyle (decreased energy expenditure) as well as genetic and epigenetic predisposition.[4]

DEFENDING A WEIGHT SET POINT: MECHANISMS OF ENERGY HOMEOSTASIS

Normally, a person's weight remains fairly constant throughout adult life. This reflects an overall balance between energy intake and expenditure. Whereas a negative feedback mechanism of peripheral fat on food intake was initially proposed more than 50 years ago,[5] our understanding of the complex and frequently redundant mechanisms involved in energy homeostasis has greatly improved over the past decade.

Central Nervous System Control of Feeding

Humans (and most vertebrates) ingest food in discrete episodes that allow time to be spent engaged in other behaviors between meals. The frequency and caloric value of each meal determines the overall energy intake. Short-term polypeptide hormonal signals from the periphery stimulate hunger and food-seeking behavior before a meal (orexigens; eg, ghrelin), whereas others promote satiety and meal cessation during and after a meal (anorectics; eg, glucagon-like peptide-1 [GLP-1], peptide tyrosine tyrosine $[\text{PYY}]_{3-36}$ [PYY_{3-36}], cholecystokinin [CCK], oxyntomodulin [OXM], and several others).[6] Long-term signals related to total body energy stores (eg, insulin, leptin, and ghrelin) may modulate the efficacy of short-term signals to defend against deviations away from a weight set point.[7]

> *Long-term signals related to total body energy stores (eg, insulin, leptin, and ghrelin) may modulate the efficacy of short-term signals to defend against deviations away from a weight set point.*

Several areas of the brain have been identified as important in the regulation of energy balance. The arcuate nucleus (ARC) of the hypothalamus contains populations of neuropeptide Y/Agouti-related protein (NPY/AgRP)- expressing neurons that are inhibited by insulin, leptin, and PYY, and stimulated by ghrelin.[8] These neurons stimulate orexigenic neurons in the lateral hypothalamic area and prefornical area. Stimulation of the NPY/AgRP neurons results in increased feeding and weight gain.[9] Also within the ARC are neurons that coexpress proopiomelanocortin and cocaine- and amphetamine-related transcript. These cells are stimulated by leptin and inhibited by the NPY/AgRP neurons.[10] In rodent and human studies, treatment with leptin caused decreased food intake and weight loss. The ARC of the hypothalamus and the lateral hypothalamic area receive and integrate information about overall energy stores from the periphery and relay this information to higher centers such as the ventral tegmental area of the mesolimbic system, which is implicated in the hedonic experience of feeding.

Peripheral Feedback Signals

The ability to increase or decrease caloric intake in response to changes in the energy status quo relies on hormonal, nutritional, and neuronal afferent feedback from the periphery to the central nervous system (CNS). Hormonal signals from adipose tissue (adipokines), including leptin and adiponectin, provide feedback signals to the brain (adipose-brain axis), whereas feedback signals from the intestinal tract (the gut-brain axis) include polypeptide hormones (enterokines) and neuronal signals via the vagus and other autonomic afferents. Basal levels of insulin and leptin increase in proportion to increasing fat mass and provide a long-term signal of body energy stores to the brain. In contrast, basal ghrelin levels are inversely related to body energy stores. Short-term fluctuations in several gut hormones including ghrelin and insulin as well as GLP-1, PYY_{3-36}, CCK, and vagal signals regulate mealtimes.

THE HORMONAL BASIS OF METABOLIC SURGERY

There has been an exponential rise in the number of metabolic (ie, bariatric) operations being performed. This is partly due to metabolic surgery achieving the most effective and durable weight loss and improvement of comorbid diseases compared with alternative treatments such as diet and exercise[11] or currently available pharmacologic[12] or endoscopic[13] interventions. This knowledge, combined with an excellent safety profile as a result of refinement and standardization of minimally invasive techniques, has greatly increased patient and physician acceptance of metabolic surgery.[14]

Metabolic surgery comprises a heterogeneous group of operations that have been previously oversimplified as purely restrictive, purely malabsorptive, or as combinations of restrictive and malabsorptive.

> *Metabolic surgery comprises a heterogeneous group of operations that have been previously oversimplified as purely restrictive, purely malabsorptive, or as combinations of restrictive and malabsorptive.*

In fact, each operation results in multiple effects, depending on the exact alterations in the gastrointestinal anatomy. For a detailed description and diagrams of the 4 currently performed metabolic operations in the United States, please refer to the article "Surgery for Diabesity" by Dr. Arroyo *et al.* in this issue. For an example, the sleeve gastrectomy (SG) has in the past been regarded as purely restrictive, as the capacity of the stomach is greatly reduced and none of the absorptive surface of the small intestine is bypassed. However, the major source of ghrelin, an appetite-stimulating hormone, is the fundus and body of the stomach, and the reduction in ghrelin following SG likely plays an important role in decreasing appetite.

Bypass-type operations such as the Roux-en-Y gastric bypass (RYGB) and biliopancreatic diversion with or without duodenal switch (BPD/DS) are traditionally considered to result in a combination of restriction and malabsorption. As well as bypassing or removing a large part of the stomach, these operations also divert chyme away from the duodenum and proximal small bowel and instead present nutrients rapidly to the distal gut. True calorie malabsorption is unlikely with RYGB, although iron, vitamins D and B12, and calcium absorption are reduced due to the bypass of their primary sites of absorption in the duodenum.

The adjustable gastric band (AGB), which places a constrictive ring around the gastric inlet and is generally considered to be a purely restrictive operation, results in prolonged satiety by as yet incompletely understood mechanisms.[15]

It is increasingly believed that both the weight loss and the improvement of the metabolic derangements associated with obesity result not simply from the decrease in caloric intake or degree of adiposity, but from alterations in gut hormone expression. Two key observations support this hypothesis. First, following metabolic surgery, while patients are in a state of negative energy balance and are rapidly losing weight, they do not complain of feeling ravenously hungry, but rather have decreased appetite and early satiety.[16] This suggests a resetting of their defended weight set point to a lower level. Second, following several types of metabolic surgery, there is a rapid improvement in glycemic control in patients with T2DM, occurring almost immediately after surgery before significant weight loss has occurred. Here we will review our current understanding of the roles of 6 hormones that are considered to be important in energy homeostasis and whose expression is altered by metabolic surgery (Table 1).

Ghrelin

Ghrelin is a 28 amino acid peptide that is produced at multiple sites of the diffuse endothelial system and brain. It was initially discovered in 1999 as the endogenous ligand for the growth hormone secretagogue receptor (GHS-R1a) in the pituitary.[17] Later, it was found that treatment with ghrelin caused an increase in feeding and weight gain in rodents

Table 1. *Anatomic and Physiologic Alterations Due to Metabolic Operations.*

	Stomach	Duodenum	Proximal Intestine	Distal Intestine
RYGB	Decreased capacity to 15–30 mL Chyme bypasses 90% of stomach Decreased gastric transit time	Bypassed by chyme	Proximal jejunum bypassed by chyme Rapid exposure of distal jejunum to chyme Bile mixes with chyme in proximal ileum	Earlier exposure to chyme
AGB	Restrictive band around gastric cardia Direct pressure on vagus nerves	No alterations	No alterations	No alterations
BPD/DS	Decreased capacity to 60–150 mL 85%–90% of stomach removed Accelerated gastric emptying by 2-fold to 3-fold	Bypassed by chyme	Chyme bypasses jejunum Rapid exposure of ileum to chyme Bile mixes with chyme in distal ileum	Rapid exposure to chyme
SG	Decreased capacity to 60–150 mL 85%–90% of stomach removed Accelerated gastric emptying by 2-fold to 3-fold	More rapid exposure	Possibly earlier exposure to chyme	Possibly earlier exposure to chyme

Abbreviations: AGB, adjustable gastric band; BPD/DS, biliopancreatic diversion with or without duodenal switch; RYGB, Roux-en-Y gastric bypass; SG, sleeve gastrectomy.

and its orexigenic role in energy balance has now been extensively studied.[9]

The greatest production of circulating ghrelin is by the ghrelin cells (aka P/D1 cells) located in the oxyntic glands of the fundus and body of the stomach. Ghrelin levels demonstrate an ultradian rhythm–rising during the day and peaking before meals.[18] Recently it has been shown that ghrelin cells express the clock genes *PER1* and *PER2* and ghrelin release is entrained to peak just prior to normally available mealtimes.[19] A rising level of ghrelin prompts an increase in food seeking behavior and the initiation of feeding. After feeding, ghrelin levels are rapidly suppressed.[18] Part of ghrelin's effect on the brain is mediated by the vagus nerves and part by a humoral effect.

Ghrelin receptors or ghrelin-stimulated activity has been demonstrated in multiple sites of the CNS understood to be important in energy homeostasis. These include the ARC, the nucleus accumbens, the parabrachial nucleus, dorsomedial and lateral hypothalamic area and the subfornical part of the circumventricular organs. In the ARC, ghrelin's orexigenic effect is mediated by stimulation of NPY/AgRP neurons.[9]

Ghrelin levels are lower in obese subjects' and higher in lean subjects. Diet-induced or exercise-induced weight loss causes a rise in fasting levels of ghrelin, and this may result in increased sensation of hunger and failure to lose excess weight.[20] Thus, ghrelin may be important in maintaining a weight set point, particularly in resisting weight loss.

Twenty-three studies report ghrelin levels before and after RYGB (Table 2). A significant early decrease in ghrelin was found in 5 of the 7 studies that reported ghrelin levels within the first 14 days after surgery. One of these studies compared patients undergoing RYGB with a control group of patients undergoing other nonresective gastric operations such as AGB or Nissen fundoplication.[21,22] Only the RYGB resulted in a significant reduction in postoperative ghrelin levels.

Reports of longer-term changes in ghrelin levels after RYGB are inconsistent in the literature (Table 2). Twenty-one studies reported preoperative and postoperative circulating ghrelin levels at 34 time intervals between 1 month and 2 years after surgery. Many of these studies measured a single ghrelin level in the morning following an overnight fast, whereas others also assessed the meal-associated suppression of ghrelin. Some studies measured "active" ghrelin, the metabolically active isoform consisting of a

Table 2. *Ghrelin Changes After Metabolic Surgery.*

Author, Year	No. of Subjects	Postoperative Interval	Fasting Ghrelin	Oral Glucose- or Meal-Suppressed Ghrelin (Nadir or AUC)
RYGB				
Lin *et al.* 2004[58]	34	30 minutes	Decreased	
Couce *et al.* 2006[59]	49	2 hours	Decreased	
Frühbeck *et al.* 2004[22]	15	1 day	Decreased	
Le Roux *et al.* 2007[36]	16	2 days	NS	
Le Roux *et al.* 2007[36]	16	4 days	NS	
Le Roux *et al.* 2007[36]	16	6 days	NS	
Sundbom *et al.* 2007[60]	15	6 days	Decreased	
Peterli *et al.* 2009[26]	13	1 week	Decreased	
Couce *et al.* 2006[59]	18	10 days	Decreased	
Korner *et al.* 2009[25]	28	14 days	NS	
Borg *et al.* 2006[61]	6	1 month	NS	
Bose *et al.* 2010[24]	11	1 month	NS	NS
Karamanakos *et al.* 2008[45]	16	1 month	NS	
Oliván *et al.* 2009[23]	11	1 month	NS	NS
Sundbom *et al.* 2007[60]	15	1 month	NS	
Le Roux *et al.* 2007[36]	16	6 weeks	NS	
Morínigo *et al.* 2004[62]	8	6 weeks	Decreased	
Morínigo *et al.* 2008[63]	25	6 weeks	Decreased	
Borg *et al.* 2006[61]	6	3 months	NS	
Karamanakos *et al.* 2008[45]	16	3 months	NS	
Korner *et al.* 2009[25]	28	3 months	NS	
Peterli *et al.* 2009[26]	13	3 months	Decreased	
Borg *et al.* 2006[61]	6	6 months	NS	
Couce *et al.* 2006[59]	11	6 months	NS	
Frühbeck *et al.* 2004[64]	8	6 months	Decreased	
Holdstock *et al.* 2003[65]	66	6 months	Increased	
Karamanakos *et al.* 2008[45]	16	6 months	Increased	
Korner *et al.* 2009[25]	28	6 months	NS	NS
Sundbom *et al.* 2007[60]	15	6 months	NS	
Vendrell *et al.* 2004[66]	34	6 months	Increased	
Whitson *et al.* 2007[67]	10	6 months	NS	
Garcia-Fuentes *et al.* 2008[68]	13	7 months	Increased	
García de la Torre 2008[69]	17	10.5 months	Decreased	
Bose *et al.* 2010[24]	11	1 year	NS	NS nadir, increased AUC
Holdstock *et al.* 2003[65]	66	1 year	Increased	
Karamanakos *et al.* 2008[45]	16	1 year	NS	
Korner *et al.* 2009[25]	28	1 year	NS	NS
Morínigo *et al.* 2008[63]	25	1 year	NS	
Pardina *et al.* 2009[70]	34	1 year	Increased	
Sundbom *et al.* 2007[60]	15	1 year	Increased	
Faraj *et al.* 2003[71]	25	12.3 months	Increased	
Faraj *et al.* 2003[71]	25	17.5 months	NS	
Roth *et al.* 2009[72]	18	23 months	Decreased	
Stoeckli *et al.* 2004[73]	5	2 years	NS	
AGB				
Frühbeck *et al.* 2004[22]	12	1 day	Increased	
Langer *et al.* 2005[27]	10	1 day	NS	
Korner *et al.* 2009[25]	15	2 weeks	NS	
Bose *et al.* 2010[24]	9	1 month	NS	NS
Langer *et al.* 2005[27]	10	1 month	Increased	
Korner *et al.* 2009[25]	15	3 months	NS	
Carroll *et al.* 2009[74]	34	6 months	Increased	
Frühbeck *et al.* 2004[64]	8	6 months	Increased	

(*continued overleaf*)

Table 2. *(Continued).*

Author, Year	No. of Subjects	Postoperative Interval	Fasting Ghrelin	Oral Glucose- or Meal-Suppressed Ghrelin (Nadir or AUC)
Hanusch-Enserer *et al.* 2003[75]	12	6 months	NS	
Hanusch-Enserer *et al.* 2004[76]	18	6 months	NS	
Korner *et al.* 2009[25]	15	6 months	NS	NS
Langer *et al.* 2005[27]	10	6 months	Increased	
Schindler *et al.* 2004[77]	23	6 months	Increased	
Shak *et al.* 2008[78]	24	6 months	NS	
Bose *et al.* 2010[24]	9	1 year	NS	NS
Hanusch-Enserer *et al.* 2003[75]	12	1 year	NS	
Hanusch-Enserer *et al.* 2004[76]	18	1 year	Increased	
Korner *et al.* 2009[25]	15	1 year	NS	Increased AUC
Shak *et al.* 2008[78]	24	1 year	NS	
Cohen *et al.* 2005[28]	15	18 months	Increased	
Stoeckli *et al.* 2004[73]	8	2 years	Increased	
Wang *et al.* 2009[79]	15	2 years	Increased	
BPD/DS				
Adami *et al.* 2003[80]	15	5 days	Decreased	
Mingrone *et al.* 2006, 2009[81,82]	6	2 weeks	Increased	
García-Unzueta *et al.* 2005[83]	30	1 month	NS	
Adami *et al.* 2003[80]	15	2 months	NS	
Adami *et al.* 2004[84]	24	2 months	NS	
García-Unzueta *et al.* 2005[83]	30	3 months	Increased	
Garcia-Fuentes *et al.* 2008[68]	38	7 months	NS	
García de la Torre 2008[69]	11	10.5 months	Decreased	
Adami *et al.* 2004[84]	24	1 year	Increased	
García-Unzueta *et al.* 2005[83]	30	1 year	Increased	
Kotidis *et al.* 2006[54]	13	18 months	Decreased	
Valera Mora *et al.* 2007[85]	11	18 months	NS	
SG				
Langer *et al.* 2005[27]	10	1 day	Decreased	
Peterli *et al.* 2009[26]	14	1 week	Decreased	
Karamanakos *et al.* 2008[45]	16	1 month	Decreased	
Langer *et al.* 2005[27]	10	1 month	Decreased	
Karamanakos *et al.* 2008[45]	16	3 months	Decreased	
Peterli *et al.* 2009[26]	14	3 months	Decreased	
Bohdjalian *et al.* 2010[86]	12	6 months	Decreased	
Cohen *et al.* 2005[28]	4	6 months	Decreased	
Karamanakos *et al.* 2008[45]	16	6 months	Decreased	
Langer *et al.* 2005[27]	10	6 months	Decreased	
Bohdjalian *et al.* 2010[86]	12	1 year	Decreased	
Karamanakos *et al.* 2008[45]	16	1 year	Decreased	
Wang *et al.* 2009[79]	10	2 years	Decreased	
Bohdjalian *et al.* 2010[86]	12	5 years	Decreased	

Abbreviations: AGB, adjustable gastric band; AUC, area under curve; BPD/DS, biliopancreatic diversion with or without duodenal switch; NS, not significant; RYGB, Roux-en-Y gastric bypass; SG, sleeve gastrectomy.

Some studies have measured "active" ghrelin, the metabolically active isoform consisting of a unique post-translational octanoylation of the serine-3 hydroxyl group,[17] whereas others have used assays that also detect the inactive des-acyl isoform.

unique post-translational octanoylation of the serine-3 hydroxyl group,[17] whereas others used assays that also detect the inactive des-acyl isoform.

This variability between studies as well as differences between assays make comparisons between studies of questionable value.

Although considerable physiologic variations in ghrelin levels may contribute to the disparate findings, 2 observations would tend to exaggerate an expected decreased level of ghrelin in patients after RYGB. First, in a landmark cross-sectional study in 2002, Cummings *et al.* measured 24-hour ghrelin profiles in 5 subjects who had lost weight after RYGB and compared them with normal-weight individuals and BMI-matched obese individuals who had lost weight by dieting alone. They found that the RYGB subjects had markedly reduced ghrelin levels overall and had also abolished the usual preprandial spike in ghrelin. Thus, both the reduced ghrelin variability and the preprandial sample timing after an overnight fast would tend to highlight a postoperative decrease in ghrelin if present. However, the findings by Cummings *et al.* have not been consistently replicated by other investigators, several studies finding an unchanged ghrelin spike before a meal and suppression following a meal after RYGB.[23–25] These inconsistencies may be due to varying techniques for performing RYGB.

Similar to the RYGB studies, changes in ghrelin levels following BPD/DS have been mixed. Eight studies measuring ghrelin before and at 12 time points following BPD/DS found significantly reduced ghrelin at 3 time points, no significant change at 5, and increased ghrelin at 4.

Nine published articles have reported on ghrelin levels before and after AGB. These have been more consistent with increases in ghrelin reported at 7 of the 17 time points and no studies finding a decrease. Recently 2 of the authors prospectively compared the hormonal changes following RYGB and AGB. Although they did not find any significant differences in either fasting ghrelin or suppression of ghrelin following a mixed liquid meal at any time point up to 1 year, there was a significant increase in the ghrelin area under the curve (AUC) at 1 year following the AGB and not the RYGB.[25]

Five studies have evaluated the changes in ghrelin following SG. Perhaps unsurprisingly for an operation that involves removal of the majority of the ghrelin-producing stomach, fasting ghrelin levels are decreased at all time points up to 5 years of follow up. Furthermore, when compared with RYGB[26] or AGB,[27,28] the SG results in decreased fasting ghrelin levels.

> *Five studies have evaluated the changes in ghrelin following sleeve gastrectomy. When compared with Roux-en-Y gastric bypass (RYGB) or adjustable gastric band, sleeve gastrectomy results in decreased fasting ghrelin levels.*

Glucagon-like Peptide-1

GLP-1 is a member of the enteroglucagon family of polypeptide hormones that include glycentin, enteroglucagon, GLP-2, OXM, and others. They, along with glucagon, are all derived from post-translational processing of the preproglucagon gene product. The truncated active circulating isoforms include GLP-1 (7–36) amide and GLP-1 (7–37).[29] Its half-life in the blood is 2 minutes, being renally cleared and inactivated by the enzyme dipeptidyl peptidase-4 (DPP-IV).

GLP-1 is secreted along with PYY and OXM by the L-cells of the distal intestine in response to a meal. It is also a neurotransmitter present in several areas of the brain, including the paraventricular nucleus and dorsomedial hypothalamus, dorsovagal complex, thalamus, and pituitary. The GLP-1 receptor is G-protein-coupled and is present in multiple sites including the pancreas, stomach, and brain.[30]

Both GLP-1 and PYY$_{3-36}$ inhibit feeding when peripherally administered to both rodent and human subjects and have been found to have additive effects on reducing food intake.[31] Together, and in combination with other anorectic gut hormones such as CCK and OXM, they signal adequate caloric delivery to the intestine and promote satiety.

> *Both glucagon-like peptide-1 (GLP-1) peptide tyrosine tyrosine$_{3-36}$ inhibit feeding when peripherally administered to both rodent and human subjects and have been found to have additive effects on reducing food intake.*

GLP-1 and PYY$_{3-36}$ act directly on the stomach to slow gastric emptying and acid secretion. In this way they act as an "ileal brake" to intestinal motility, preventing too-rapid transit and allowing time for digestion and absorption of nutrients in the proximal intestine.[32]

GLP-1, one of the gut hormones termed incretins, has an important role in enhancing insulin secretion by the pancreas. The "incretin effect" is the increased insulin response to enteral as opposed to parenteral glucose.[33] Both in vivo and in vitro studies have demonstrated increased insulin secretion by the beta cell in response to GLP-1. Furthermore, GLP-1 inhibits glucagon secretion and increases insulin sensitivity. GLP-1 protects against apoptosis and promotes beta cell regeneration.[34] Diabetic patients demonstrate a blunted first-phase insulin response following a meal partly due to a decreased incretin response.[35] Exenatide, a GLP-1 receptor agonist, is approved for the treatment of T2DM.

The rapid resolution of diabetes following RYGB may be due to altered expression of the incretins, such as GLP-1. Following RYGB, fasting GLP-1 levels are generally unchanged (Table 3). However, the GLP-1 response to a glucose or mixed meal challenge is consistently enhanced. This effect occurs within 2 days of surgery[36] and remains at 1 year.[24,25,37]

The rapid resolution of diabetes following RYGB may be due to altered expression of the incretins, such as GLP-1.

Following AGB, the GLP-1 response to a glycemic challenge is unchanged and is significantly lower compared with RYGB.[24,25] Diabetes resolution tends to occur more gradually and is associated with weight loss.[38]

Six studies have looked at GLP-1 following BPD/DS. Similar to RYGB, all have shown unchanged fasting GLP-1 and a significantly augmented response to a glycemic challenge.

Only a single study has prospectively studied GLP-1 changes following SG. Peterli *et al.* randomized patients to receive either SG or RYGB.[26] Interestingly, following a test meal, the SG patients had an augmented GLP-1 and PYY response similar to the RYGB patients. In addition, there was a rapid resolution of diabetes in both groups. Their results do not support the foregut theory that bypass of the duodenum and proximal small intestine is important in diabetes remission following bypass operations. Another explanation is that the rapid gastric transit that has been demonstrated following SG results in early exposure of the distal intestine to nutrients and stimulation of the "ileal brake" mechanism. Ghrelin levels were significantly decreased after both SG and RYGB, and the diabetogenic and orexigenic role of ghrelin may also play a role in diabetes remission.

Peptide Tyrosine Tyrosine

PYY_{1-36} is cosecreted with GLP-1 and OXM by L-cells of the distal small bowel and colon in proportion to caloric content of a meal. The main circulating and active form, PYY_{3-36}, is formed by cleavage of PYY_{1-36} by the enzyme DPP-IV.[39] Levels rise within 15 minutes of initiating a meal and peak 1–2 hours later, remaining elevated for several hours.[40] An anorectic effect of PYY_{3-36} was first reported by Batterham *et al.* in 2002.[41] Similar to GLP-1, it has an anorectic action on the CNS, reduces gastric emptying and acid production, inhibits pancreatic exocrine function, and slows intestinal transit time. A reduced PYY_{3-36} response to feeding has been identified in obese subjects and may represent one cause of excessive appetite.[42]

The anorectic action of PYY_{3-36} on the ARC is via the Y2 receptors which are expressed on NPY/AgRP neurons. Y2 receptors are also expressed in the nucleus of the solitary tract and the nodose ganglion of the vagus, and most of PYY_{3-36} anorectic effect is likely mediated by the vagus nerve.[43]

Similar to GLP-1, following RYGB there is an early and dramatic increase in PYY_{3-36} response to a glycemic challenge (Table 4). Along with GLP-1, this contributes to the decreased appetite and rapid weight loss.[39,44] Three studies also reported a significantly increased fasting level of PYY_{3-36} following RYGB.

Two studies compared the meal-stimulated PYY_{3-36} response of RYGB and AGB.[24,25] Neither demonstrated a change in either the fasting PYY or the meal-stimulated response following the AGB. This was in contrast to the enhanced response following RYGB.

A single prospective study found an increased fasting PYY_{3-36} before and 7 months following the BPD/DS.

Three studies have evaluated changes in the meal-stimulated PYY_{3-36} response before and after SG.[26,45,46] Further supporting the findings of Peterli *et al.* with respect to an increased GLP-1 response to SG, all 3 studies found an increased meal-stimulated PYY_{3-36} response. This was similar to the change in PYY_{3-36} after RYGB.

Taken together, the GLP-1 and PYY_{3-36} changes are similar after RYGB, SG, and BPD/DS. The AGB, however, does not appear to alter the expression of these hormones.

Insulin

Insulin plays several key roles in energy metabolism and has been more extensively studied than any other hormone. Release of insulin by pancreatic beta cells is stimulated by multiple factors, in particular the elevated glucose levels following a meal. Other factors stimulating insulin release include several amino acids (alanine, glycine, and arginine), acetylcholine from vagal nerve endings, and certain incretins, including GLP-1 and glucose-dependent insulinotropic polypeptide (GIP).

Table 3. *GLP-1 Changes After Metabolic Surgery.*

Author, Year	No. of Subjects	Postoperative Interval	Fasting GLP-1	Oral Glucose- or Meal-Stimulated GLP-1 (Peak or AUC)
RYGB				
Le Roux *et al.* 2007[36]	16	2 days	NS	Increased
Le Roux *et al.* 2007[36]	16	4 days	NS	Increased
Le Roux *et al.* 2007[36]	16	6 days	NS	Increased
Kashyap *et al.* 2010[87]	9 T2DM	1 week	NS	Increased
Peterli *et al.* 2009[26]	13	1 week	NS	Increased
Clements *et al.* 2004[88]	20	2 weeks	NS	
Rubino *et al.* 2004[89]	10	3 weeks	NS	
Borg *et al.* 2006[61]	6	1 month	NS	Increased
Bose *et al.* 2010[24]	11	1 month	NS	Increased
Kashyap *et al.* 2010[87]	9 T2DM	1 month	NS	Increased
Laferrère *et al.* 2007[90]	8 T2DM	1 month	NS	Increased
Oliván *et al.* 2009[23]	11	1 month	NS	Increased
Clements *et al.* 2004[88]	20	6 weeks	NS	
Le Roux *et al.* 2007[36]	16	6 weeks	NS	Increased
Morínigo *et al.* 2006[37]	12 NGT, 12 IGT, 10 T2DM	6 weeks	NS	Increased (NS in T2DM group)
Borg *et al.* 2006[61]	6	3 months	NS	Increased
Clements *et al.* 2004[88]	20	3 months	NS	
Peterli *et al.* 2009[26]	13	3 months	NS	Increased
Borg *et al.* 2006[61]	6	6 months	NS	Increased
Korner *et al.* 2009[25]	28	6 months	NS	Increased
Whitson *et al.* 2007[67]	5 T2DM, 5 NGT	6 months	NS	
Bose *et al.* 2010[24]	11	1 year	NS	Increased
Korner *et al.* 2009[25]	28	1 year	NS	Increased
Morínigo *et al.* 2006[37]	12 NGT, 12 IGT, 10 T2DM	1 year	NS	Increased
Pournaras *et al.* 2010[91]	6	18 months	NS	
Reinehr *et al.* 2007[92]	19	2 years	Decreased	
De Carvalho *et al.* 2009[93]	19 (8 T2DM)		NS	Increased
AGB				
Bose *et al.* 2010[24]	9	1 month	NS	NS
Carroll *et al.* 2009[74]	34	6 months	Increased	
Korner *et al.* 2009[25]	15	6 months	NS	NS
Shak *et al.* 2008[78]	24	6 months	NS	
Bose *et al.* 2010[24]	9	1 year	NS	NS
Korner *et al.* 2009[25]	15	1 year	NS	NS
Shak *et al.* 2008[78]	24	1 year	NS	
Reinehr *et al.* 2007[92]	11	2 years	Decreased	
BPD/DS				
Guidone *et al.* 2006[94]	10 T2DM	0.25	NS	Increased
Mingrone *et al.* 2009[82]	10 (5 T2DM)	0.5	NS	Increased
Guidone *et al.* 2006[94]	10 T2DM	1	NS	Increased
Marfella *et al.* 2010[95]	36 T2DM	1	NS	Increased
Salinari *et al.* 2009[96]	9 T2DM	1	NS	Increased
Valverde *et al.* 2005[97]	19	1	NS	Increased
Valverde *et al.* 2005[97]	19	3	Increased	Increased
Valverde *et al.* 2005[97]	19	6	Increased	Increased
Lugari *et al.* 2004[98]	22 (non-T2DM)	8.7	Increased	Increased
SG				
Peterli *et al.* 2009[26]	14	1 week	NS	Increased
Peterli *et al.* 2009[26]	14	3 months	NS	Increased

Abbreviations: AGB, adjustable gastric band; AUC, area under curve; BPD/DS, biliopancreatic diversion with or without duodenal switch; GLP-1, glucagon-like peptide-1; IGT, impaired glucose tolerance; NGT, normal glucose tolerance; NS, not significant; RYGB, Roux-en-Y gastric bypass; SG, sleeve gastrectomy; T2DM, type 2 diabetes mellitus.

Table 4. *PYY Changes After Metabolic Surgery.*

Author, Year	No. of Subjects	Postoperative Interval	Fasting PYY	Oral Glucose- or Meal-Stimulated PYY (Peak or AUC)
RYGB				
Le Roux *et al.* 2007[36]	16	2 days		Increased
Le Roux *et al.* 2007[36]	16	4 days		Increased
Le Roux *et al.* 2007[36]	16	6 days		NS
Peterli *et al.* 2009[26]	13	1 week		Increased
Korner *et al.* 2009[25]	28	2 weeks	NS	
Karamanakos *et al.* 2008[45]	16	1 month	Increased	Increased
Bose *et al.* 2010[24]	11	1 month	NS	Increased
Borg *et al.* 2006[61]	6	1 month	NS	Increased
Oliván *et al.* 2009[23]	11	1 month	Decreased	Increased
Morínigo *et al.* 2008[63]	25	6 weeks	NS	
Le Roux *et al.* 2007[36]	16	6 weeks		Increased
Valderas *et al.* 2010[46]	8	2 months	NS	Increased
Peterli *et al.* 2009[26]	13	3 months		Increased
Korner *et al.* 2009[25]	28	3 months	NS	
Karamanakos *et al.* 2008[45]	16	3 months	Increased	Increased
Borg *et al.* 2006[61]	6	3 months	NS	Increased
Korner *et al.* 2009[25]	28	6 months	NS	Increased
Karamanakos *et al.* 2008[45]	16	6 months	Increased	Increased
Borg *et al.* 2006[61]	6	6 months	NS	Increased
Garcia-Fuentes *et al.* 2008[68]	13	7 months	Increased	
Morínigo *et al.* 2008[63]	25	1 year	NS	
Korner *et al.* 2009[25]	28	1 year	Increased	Increased
Karamanakos *et al.* 2008[45]	16	1 year	Increased	Increased
Bose *et al.* 2010[24]	11	1 year	NS	Increased
Pournaras *et al.* 2010[91]	6	18–24 months		Increased
Reinehr *et al.* 2007[92]	19	2 years	Increased	
AGB				
Korner *et al.* 2009[25]	15	2 weeks	NS	
Bose *et al.* 2010[24]	9	1 month	NS	NS
Korner *et al.* 2009[25]	15	3 months	NS	
Korner *et al.* 2009[25]	15	6 months	NS	Increased
Korner *et al.* 2009[25]	15	1 year	NS	NS
Bose *et al.* 2010[24]	9	1 year	NS	NS
Reinehr *et al.* 2007[92]	11	2 years	Increased	
BPD/DS				
Garcia-Fuentes *et al.* 2008[68]	38	7 months	Increased	
SG				
Peterli *et al.* 2009[26]	14	1 week		Increased
Karamanakos *et al.* 2008[45]	16	1 month	Increased	
Valderas *et al.* 2010[46]	8	2 months	NS	Increased
Karamanakos *et al.* 2008[45]	16	3 months	Increased	Increased
Peterli *et al.* 2009[26]	14	3 months		Increased
Karamanakos *et al.* 2008[45]	16	6 months	Increased	Increased
Karamanakos *et al.* 2008[45]	16	1 year	Increased	Increased

Abbreviations: AGB, adjustable gastric band; AUC, area under curve; BPD/DS, biliopancreatic diversion with or without duodenal switch; NS, not significant; PYY, peptide tyrosine-tyrosine; RYGB, Roux-en-Y gastric bypass; SG, sleeve gastrectomy.

Insulin has multiple physiologic actions related to energy balance and its vital importance is demonstrated by the high degree of homology of the insulin molecule between vertebrate and even some invertebrate species. Postprandial spikes in insulin serve to maintain normal glucose levels following a meal by promoting energy storage as glycogen in muscle and liver and as triglycerides in adipose tissue. A low level of insulin in times of fasting reverses these actions and results in mobilization of glucose from the liver and muscle and mobilization of energy stores from adipose tissue.

Insulin has both long-term and short-term effects on energy homeostasis. Insulin appears to have a negative feedback role on feeding by direct effects on the CNS. In animal studies, when injected centrally, insulin causes decreased food intake and weight loss. Insulin inhibits NPY/AgRP neurons in the ARC. Insulin levels have also been suggested as an important long-term indicator of overall energy stores. Higher basal and meal-stimulated insulin levels are demonstrated in overweight compared with lean subjects.[47]

> *Insulin has both long-term and short-term effects on energy homeostasis.*

Following RYGB and BPD/DS, fasting insulin levels are decreased (Table 5). In obese patients with T2DM or impaired glucose tolerance there is a rapid improvement in glycemic control. This is due to both a restored first-phase insulin response to a glycemic challenge due to an enhanced incretin effect, and to improved insulin sensitivity related to weight loss.

Fasting insulin levels are also reduced after AGB. However, the enhanced incretin effect in response to a meal is not seen. In diabetic patients, insulin sensitivity improves with weight loss.

Interestingly, SG also results in a rapid resolution of diabetes in T2DM patients with decreased fasting insulin levels and a rapid improvement of the first phase insulin response and insulin sensitivity.[26,48]

Leptin

Leptin is a 167 amino acid protein produced at several sites in the body, the majority being produced by adipocytes and it acts as a signal of overall body energy stores rising in proportion to total body fat mass. It was discovered in 1994 as a result of work on mice that expressed a Mendelian recessively inherited enhanced appetite and extreme obesity, the ob/ob phenotype.[49]

Leptin is produced mainly in white adipose tissue, although it is also produced in brown adipose tissue and the placenta, ovaries, skeletal muscle, stomach, breast, bone marrow, pituitary, and liver.

Leptin acts on the hypothalamus, where it inhibits NPY/AgRP receptor neurons in the ARC while stimulating αMSH neurons.[8,10] Leptin binds to the OB-Rb receptor and activates the JAK-STAT and MAPK signal transduction pathways.

Its effects on appetite and energy balance are largely opposite to the action of ghrelin. Although in mice leptin appears to act as an important physiological brake on appetite, its role in humans has come into question due to the minimal weight loss seen in obese patients taking leptin analogues.

> *Although in mice leptin appears to act as an important physiological brake on appetite, its role in humans has come into question due to the minimal weight loss seen in obese patients taking leptin analogues.*

Leptin resistance in obesity has been proposed as one explanation of this finding. Another hypothesis is that, in humans, relative leptin deficiency is actually the more potent stimulus at times of negative energy balance, such as starvation and diet-induced weight loss.

Studies measuring leptin after metabolic surgery consistently demonstrate a decrease with weight loss (Table 6). Oliván *et al.* did not find a significant difference between leptin levels after equivalent weight loss after either RYGB or diet/exercise.[23]

Adiponectin

High-molecular-weight adiponectin is a 244 amino acid protein produced in white adipose tissue. Serum adiponectin levels are inversely proportional to body fat mass, BMI, waist-to-hip ratio, serum insulin, and glucose levels. In a study by English *et al.*, fasting adiponectin levels were decreased in obese individuals compared with lean controls. However, the obese subjects had an exaggerated response to a meal.[50]

Adiponectin is suspected to have a role in the pathogenesis of insulin resistance in T2DM. In rodents, adiponectin administration improves hepatic gluconeogenesis and increases fatty acid utilization in muscle cells.[51] Furthermore, in a study comparing Pima Indians who develop diabetes due to their genetic makeup and Caucasians, adiponectin levels correlated closely with insulin resistance,[52] whereas high levels are protective against developing T2DM.[53]

In 12 studies documenting adiponectin levels before and after RYGB, there was an increase in fasting adiponectin seen in all but 1 study (Table 7). After BPD/DS, 3 of 4 studies did not find a significant difference in adiponectin despite a decrease of 44% total body weight in the Kotidis *et al.* study.[54]

Following AGB, adiponectin levels also rise as weight reduces. No study looked at the changes in adiponectin following SG.

Table 5. *Insulin Changes After Metabolic Surgery.*

Author, Year	No. of Subjects	Postoperative Interval	Fasting Insulin	Oral Glucose- or Meal-Stimulated Insulin (Peak or AUC)
RYGB				
Couce *et al.* 2006[59]	49 (20 T2DM)	2 hours	Increased	
Frühbeck *et al.* 2004[22]	15	1 day	Decreased	
Kashyap *et al.* 2010[87]	9 T2DM	1 week	Decreased	
Peterli *et al.* 2009[26]	13	1 week	Decreased	Increased
Couce *et al.* 2006[59]	18	10 days	Decreased	
Clements *et al.* 2004[88]	20	2 weeks	NS	
Korner *et al.* 2009[25]	28	2 weeks	NS	
Rubino *et al.* 2004[89]	10	3 weeks	Decreased	
Borg *et al.* 2006[61]	6	1 month	Decreased	
Kashyap *et al.* 2010[87]	9 T2DM	1 month	Decreased	
Laferrère *et al.* 2007[90]	8 T2DM	1 month	Decreased	Increased peak, NS AUC
Molina *et al.* 2003[99]	29	1 month	Decreased	
Oliván *et al.* 2009[23]	11	1 month	Decreased	Increased peak
Ramos *et al.* 2006[100]	30	1 month	Decreased	
Swarbrick *et al.* 2008[101]	19	1 month	Decreased	
Clements *et al.* 2004[88]	20	6 weeks	NS	
Morínigo *et al.* 2006[37]	12 NGT, 12 IGT, 12 T2DM	6 weeks		Increased AUC
Morínigo *et al.* 2004[62]	8	6 weeks	Decreased	
Borg *et al.* 2006[61]	6	3 months	Decreased	
Clements *et al.* 2004[88]	20	3 months	NS	
Korner *et al.* 2009[25]	28	3 months	Decreased	
Molina *et al.* 2003[99]	29	3 months	Decreased	
Peterli *et al.* 2009[26]	13	3 months	Decreased	Increased peak
Ramos *et al.* 2006[100]	30	3 months	Decreased	
Swarbrick *et al.* 2008[101]	19	3 months	Decreased	
Borg *et al.* 2006[61]	6	6 months	Decreased	
Couce *et al.* 2006[59]	11	6 months	Decreased	
Frühbeck *et al.* 2004[64]	8	6 months	Decreased	
Holdstock *et al.* 2003[65]	66	6 months	Decreased	
Korner *et al.* 2009[25]	28	6 months	Decreased	NS peak
Lin *et al.* 2007[58]	28	6 months	Decreased	
Molina *et al.* 2003[99]	29	6 months	Decreased	
Swarbrick *et al.* 2008[101]	19	6 months	Decreased	
Vendrell *et al.* 2004[66]	34	6 months	Decreased	
Whitson *et al.* 2007[67]	10 (5 T2DM)	6 months	NS	
De Carvalho *et al.* 2009[93]	19 (11 T2DM)	9 months	Decreased	Increased peak
Trakhtenbroit *et al.* 2009[102]	10	9 months	NS	
Coughlin *et al.* 2007[103]	6	1 year	Decreased	
Holdstock *et al.* 2003[65]	66	1 year	Decreased	
Korner *et al.* 2009[25]	28	1 year	Decreased	NS peak
Morínigo *et al.* 2006[37]	12 NGT, 12 IGT, 12 T2DM	1 year		Increased AUC (except NGT group)
Riedl *et al.* 2008[104]	30	1 year	Decreased	
Serra *et al.* 2006[105]	70	1 year	Decreased	
Swarbrick *et al.* 2006[106]	19	1 year	Decreased	
Swarbrick *et al.* 2008[101]	19	1 year	Decreased	
Vila *et al.* 2009[107]	28	1 year	Decreased	
Vilarrasa *et al.* 2007[108]	65	1 year	Decreased	
García de la Torre 2008[69]	17	9–12 months	Decreased	
Faraj *et al.* 2003[71]	17	12.3 months	NS	
Faraj *et al.* 2003[71]	13	17.5 months	NS	
Roth *et al.* 2009[72]	18	23 months	Decreased	Decreased AUC
Reinehr *et al.* 2007[92]	19	2 years	Decreased	
García de la Torre 2008[69]	17	9–12 months	Decreased	
Pardina *et al.* 2009[70]	34	1, 3, 6, 9, 12 months	Decreased	

(*continued overleaf*)

Table 5. *(Continued).*

Author, Year	No. of Subjects	Postoperative Interval	Fasting Insulin	Oral Glucose- or Meal-Stimulated Insulin (Peak or AUC)
AGB				
Frühbeck *et al.* 2002[21]	7	1 day	Decreased	
Frühbeck *et al.* 2004[22]	12	1 day	Decreased	
Korner *et al.* 2009[25]	15	2 weeks	NS	
Korner *et al.* 2009[25]	15	3 months	NS	
Uzun *et al.* 2004[109]	40	6 months	Decreased	
Moschen *et al.* 2009[110]	30	6 months	Decreased	
Carroll *et al.* 2009[74]	34	6 months	Decreased	
Korner *et al.* 2009[25]	15	6 months	Decreased	Decreased peak
Hanusch-Enserer *et al.* 2004[76]	18	6 months	Decreased	
Shak *et al.* 2008[78]	24	6 months	NS	
Ram/Diker *et al.* 2005/2006[111,112]	23	6 months	Decreased	
Haider *et al.* 2006[113]	31	6 months	Decreased	
Frühbeck *et al.* 2004[64]	8	6 months	Decreased	
Moschen *et al.* 2009[110]	30	1 year	Decreased	
Riedl *et al.* 2008[104]	10	1 year	Decreased	
Korner *et al.* 2009[25]	15	1 year	Decreased	Decreased peak
Hanusch-Enserer *et al.* 2004[76]	18	1 year	Decreased	
Shak *et al.* 2008[78]	24	1 year	NS	
Engl *et al.* 2007[114]	13	1 year	Decreased	
Poitou *et al.* 2005[115]	65	1 year	Decreased	
Ram/Diker *et al.* 2005/2006[111,112]	23	14 months	Decreased	
Reinehr *et al.* 2007[92]	11	2 years	Decreased	
BPD/DS				
Guidone *et al.* 2006[94]	10	1 week	Decreased	Decreased AUC
Salinari *et al.* 2008[116]	9 T2DM	8 days	Decreased	Decreased AUC
Mingrone *et al.* 2009[82]	10 (5 T2DM)	2 weeks		Decreased AUC
Salinari *et al.* 2009[96]	9 T2DM	1 month		Decreased AUC
Marfella *et al.* 2010[95]	36 T2DM	1 month	Decreased	Decreased AUC
Guidone *et al.* 2006[94]	10	1 month	Decreased	Decreased AUC
Camastra *et al.* 2009[117]	16	6 months	Decreased	Decreased 24-hour insulin
Lugari *et al.* 2004[98]	22 non-T2DM	8.7 months	Decreased	Decreased AUC
García de la Torre 2008[69]	11	9–12 months	Decreased	
Valera Mora *et al.* 2007[85]	11	18 months	Decreased	
Rosa *et al.* 2007[118]	8	3 years	Decreased	Decreased 24-hour insulin
SG				
Peterli *et al.* 2009[26]	14	1 week	Decreased	Increased peak
Lee *et al.* 2010[48]	20	1 month	Decreased	Decreased AUC
Peterli *et al.* 2009[26]	14	3 months	Decreased	Increased peak
Lee *et al.* 2010[48]	20	3 months	Decreased	Decreased AUC
Lee *et al.* 2010[48]	20	6 months	Decreased	Decreased AUC
Lee *et al.* 2010[48]	20	1 year	Decreased	Decreased AUC

Abbreviations: AGB, adjustable gastric band; AUC, area under curve; BPD/DS, biliopancreatic diversion with or without duodenal switch; IGT, impaired glucose tolerance; NGT, normal glucose tolerance; NS, not significant; RYGB, Roux-en-Y gastric bypass; SG, sleeve gastrectomy; T2DM, type 2 diabetes mellitus.

TAILORED METABOLIC OPERATIONS

With our increasing understanding of the mechanisms controlling energy homeostasis, newer metabolic operations have been developed to achieve particular metabolic goals. At the current time, operations designed to treat T2DM in nonmorbidly obese patients, such as the ileal interposition[55] and duodenal-jejunal bypass,[56] are being evaluated in institutional review board–approved studies.

Table 6. *Leptin Changes After Metabolic Surgery.*

Author, Year	No. of Subjects	Postoperative Interval	Fasting Leptin
RYGB			
Korner *et al.* 2009[25]	28	2 weeks	Decreased
Rubino *et al.* 2004[89]	10	3 weeks	Decreased
Borg *et al.* 2006[61]	6	1 month	Decreased
Bose *et al.* 2010[24]	11	1 month	Decreased
Molina *et al.* 2003[99]	29	1 month	Decreased
Oliván *et al.* 2009[23]	11	1 month	Decreased
Ramos *et al.* 2006[100]	30	1 month	Decreased
Swarbrick *et al.* 2008[101]	19	1 month	Decreased
Morínigo *et al.* 2004[62]	8	6 weeks	Decreased
Borg *et al.* 2006[61]	6	3 months	Decreased
Korner *et al.* 2009[25]	28	3 months	Decreased
Molina *et al.* 2003[99]	29	3 months	Decreased
Ramos *et al.* 2006[100]	30	3 months	Decreased
Swarbrick *et al.* 2008[101]	19	3 months	Decreased
Borg *et al.* 2006[61]	6	6 months	Decreased
Bruno *et al.* 2010[119]		6 months	Decreased
Bruno *et al.* 2010[119]		6 months	Decreased
Frühbeck *et al.* 2004[64]	8	6 months	Decreased
Holdstock *et al.* 2003[65]	66	6 months	Decreased
Korner *et al.* 2009[25]	28	6 months	Decreased
Molina *et al.* 2003[99]	29	6 months	Decreased
Swarbrick *et al.* 2008[101]	19	6 months	Decreased
Vendrell *et al.* 2004[66]	34	6 months	Decreased
Garcia-Fuentes *et al.* 2008[68]	13	7 months	Decreased
García de la Torre 2008[69]	17	9–12 months	Decreased
Bose *et al.* 2010[24]	11	1 year	Decreased
Holdstock *et al.* 2003[65]	66	1 year	Decreased
Korner *et al.* 2009[25]	28	1 year	Decreased
Riedl *et al.* 2008[104]	30	1 year	Decreased
Swarbrick *et al.* 2008[101]	19	1 year	Decreased
Vila *et al.* 2009[107]	28	1 year	Decreased
Faraj *et al.* 2003[71]	25	12.3 months	Decreased
Das *et al.* 2003[120]	30	14 months	Decreased
Faraj *et al.* 2003[71]	25	17.5 months	Decreased
Czupryniak *et al.* 2007[121]	68	2 years	Decreased
Stoeckli *et al.* 2004[73]	5	2 years	Decreased
Pardina *et al.* 2009[70]	34	1, 3, 6, 9, 12	Decreased
AGB			
Frühbeck *et al.* 2002[21]	7	1 day	Increased
Maruna *et al.* 2001[122]	18	2 days	Increased
Korner *et al.* 2009[25]	15	2 weeks	Decreased
Bose *et al.* 2010[24]	9	1 month	NS
Korner *et al.* 2009[25]	15	3 months	Decreased
Carroll *et al.* 2009[74]	34	6 months	Decreased
Ram/Diker *et al.* 2005/2006[111,112]	23	6 months	Decreased
Frühbeck *et al.* 2004[64]	8	6 months	Decreased
Haider *et al.* 2006[113]	31	6 months	Decreased
Hanusch-Enserer *et al.* 2003[75]	12	6 months	Decreased
Hanusch-Enserer *et al.* 2004[76]	18	6 months	Decreased
Korner *et al.* 2009[25]	15	6 months	Decreased
Moschen *et al.* 2009[110]	30	6 months	Decreased
Shak *et al.* 2008[78]	24	6 months	Decreased
Uzun *et al.* 2004[109]	40	6 months	Decreased
Bose *et al.* 2010[24]	9	1 year	NS
Hanusch-Enserer *et al.* 2003[75]	12	1 year	Decreased
Hanusch-Enserer *et al.* 2004[76]	18	1 year	Decreased
Heinonen *et al.* 2005[123]	8	1 year	Decreased
Korner *et al.* 2009[25]	15	1 year	Decreased

(continued overleaf)

Table 6. *(Continued).*

Author, Year	No. of Subjects	Postoperative Interval	Fasting Leptin
Moschen *et al.* 2009[110]	30	1 year	Decreased
Poitou *et al.* 2005[115]	65	1 year	Decreased
Riedl *et al.* 2008[104]	10	1 year	Decreased
Shak *et al.* 2008[78]	24	1 year	Decreased
Ram/Diker *et al.* 2005/2006[111,112]	23	14 months	Decreased
Stoeckli *et al.* 2004[73]	8	2 years	Decreased
BPD/DS			
Guidone *et al.* 2006[94]	10	1 week	Decreased
García-Unzueta *et al.* 2005[83]	30	1 month	Decreased
Guidone *et al.* 2006[94]	10	1 month	Decreased
Adami *et al.* 2004[84]	24	2 months	Decreased
García-Unzueta *et al.* 2005[83]	30	3 months	Decreased
Camastra *et al.* 2009[117]	16	6 months	Decreased
Garcia-Fuentes *et al.* 2008[68]	38	7 months	NS
García de la Torre 2008[69]	11	9–12 months	Decreased
Adami *et al.* 2004[84]	24	1 year	Decreased
García-Unzueta *et al.* 2005[83]	30	1 year	Decreased
Kotidis *et al.* 2006[54]	13	18 months	NS
Valera Mora *et al.* 2007[85]	11	18 months	Decreased

Abbreviations: AGB, adjustable gastric band; BPD/DS, biliopancreatic diversion with or without duodenal switch; NS, not significant; RYGB, Roux-en-Y gastric bypass.

Table 7. *Adiponectin Changes After Metabolic Surgery.*

Author, Year	No. of Subjects	Postoperative Interval	Fasting Adiponectin
RYGB			
Couce *et al.* 2006[59]	49 (20 T2DM)	2 hours	Decreased
Couce *et al.* 2006[59]	18	10 days	NS
Bruno *et al.* 2010[119]	20	6 months	Increased
Couce *et al.* 2006[59]	11	6 months	Increased
Vendrell *et al.* 2004[66]	34	6 months	Increased
Holdstock *et al.* 2003[65]	66	6 months	Increased
Morínigo *et al.* 2004[62]	8	6 months	Increased
De Carvalho *et al.* 2009[93]	19 (8 T2DM)	9 months	Increased
García de la Torre 2008[69]	17	9–12 months	Increased
Swarbrick *et al.* 2006[106]	19	1 year	Increased
Coughlin *et al.* 2007[103]	6	1 year	Increased
Vilarrasa *et al.* 2007[108]	65	1 year	NS
Serra *et al.* 2006[105]	70	1 year	Increased
Holdstock *et al.* 2003[65]	66	1 year	Increased
Faraj *et al.* 2003[71]	25	12.3 months	Increased
Faraj *et al.* 2003[71]	25	17.5 months	Increased
AGB			
Ram/Diker *et al.* 2005/2006[111,112]	23	6 months	Increased
Haider *et al.* 2006[113]	31	6 months	Increased
Moschen *et al.* 2009[110]	30	6 months	Increased
Engl *et al.* 2007[114]	13	1 year	NS
Poitou *et al.* 2005[115]	65	1 year	Increased
Moschen *et al.* 2009[110]	30	1 year	Increased
Ram/Diker *et al.* 2005/2006[111,112]	23	14 months	Increased
BPD/DS			
Guidone *et al.* 2006[94]	10	1 week	NS
Guidone *et al.* 2006[94]	10	1 month	NS
García de la Torre 2008[69]	11	9–12 months	Increased
Kotidis *et al.* 2006[54]	13	18 months	NS

Abbreviations: AGB, adjustable gastric band; BPD/DS, biliopancreatic diversion with or without duodenal switch; NS, not significant; RYGB, Roux-en-Y gastric bypass; T2DM, type 2 diabetes mellitus.

At the current time, operations designed to treat type 2 diabetes mellitus in nonmorbidly obese patients, such as the ileal interposition and duodenal-jejunal bypass, are being evaluated in institutional review board–approved studies.

COMBINING METABOLIC SURGERY AND PHARMACOTHERAPY TO OPTIMIZE THERAPY

Not all patients have the desired therapeutic effect from their metabolic operation. For example, up to 15% of patients fail to achieve an excess weight loss of 30%[57] and a minority of diabetic patients will still have inadequate glycemic control following their metabolic operation.[14] In other patients, weight regain is accompanied by return of their presurgical conditions. A hormonal panel may point to a particular feedback system that is limiting the overall success. Targeted pharmacotherapy may be necessary and sufficient to suppress the system and achieve the metabolic goal. An example of such modulation was published in 2006. Le Roux *et al.* performed jejunoileal bypass on rats and compared these with sham-operated rats. Weight loss in the jejunoileal bypass rats was found to be due to increased PYY_{3-36}. Antagonism of the PYY with a specific neutralizing antibody caused the rats to regain the weight, whereas administration of PYY_{3-36} resulted in additional weight loss.

CONCLUSION

Metabolic surgery results in changes in the hormonal milieu that both promote weight loss in obese individuals and have dramatic beneficial effects on comorbid diseases such as T2DM and the metabolic syndrome. Continued translational research into the relative contributions of redundant signaling pathways in energy homeostasis will likely result in combinational treatment strategies including both surgery and pharmacotherapy for optimal results.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999–2008. *JAMA* 2010; 303(3): 235–241.
2. Dixon JB. The effect of obesity on health outcomes. *Mol Cell Endocrinol* 2010; 316: 104–108.
3. Poole R, Byrne CD. The metabolic syndrome and type 2 diabetes. *Minerva Endocrinol* 2005; 30: 139–159.
4. Bruce KD, Hanson MA. The developmental origins, mechanisms, and implications of metabolic syndrome. *J Nutr* 2010; 140: 648–652.
5. Kennedy GC. The role of depot fat in the hypothalamic control of food intake in the rat. *Proc R Soc Lond B Biol Sci* 1953; 140: 578–596.
6. Gibbs J, Young RC, Smith GP. Cholecystokinin decreases food intake in rats. *J Comp Physiol Psychol* 1973; 84: 488–495.
7. Konner AC, Klockener T, Bruning JC. Control of energy homeostasis by insulin and leptin: targeting the arcuate nucleus and beyond. *Physiol Behav* 2009; 97: 632–638.
8. Flier JS. Obesity wars: molecular progress confronts an expanding epidemic. *Cell* 2004; 116: 337–350.
9. Nakazato M, Murakami N, Date Y, et al. A role for ghrelin in the central regulation of feeding. *Nature* 2001; 409: 194–198.
10. Cowley MA, Smart JL, Rubinstein M, et al. Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature* 2001; 411: 480–484.
11. Picot J, Jones J, Colquitt JL, et al. The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation. *Health Technol Assess* 2009; 13(1): 190, 215–357, iii–iv.
12. Aylwin S, Al-Zaman Y. Emerging concepts in the medical and surgical treatment of obesity. *Front Horm Res* 2008; 36: 229–259.
13. Schauer P, Chand B, Brethauer S. New applications for endoscopy: the emerging field of endoluminal and transgastric bariatric surgery. *Surg Endosc* 2007; 21: 347–356.
14. Buchwald H, Estok R, Fahrbach K, et al. Weight and type 2 diabetes after bariatric surgery: systematic review and meta-analysis. *Am J Med* 2009; 122: 248–256.e5.
15. Dixon AF, Dixon JB, O'Brien PE. Laparoscopic adjustable gastric banding induces prolonged satiety: a randomized blind crossover study. *J Clin Endocrinol Metab* 2005; 90: 813–819.
16. Marceau P, Cabanac M, Frankham PC, et al. Accelerated satiation after duodenal switch. *Surg Obes Relat Dis* 2005; 1: 408–412.
17. Kojima M, Hosoda H, Date Y, et al. Ghrelin is a growth-hormone-releasing acylated peptide from stomach. *Nature* 1999; 402: 656–660.
18. Cummings DE, Purnell JQ, Frayo RS, et al. A preprandial rise in plasma ghrelin levels suggests a role in meal initiation in humans. *Diabetes* 2001; 50: 1714–1719.
19. LeSauter J, Hoque N, Weintraub M, et al. Stomach ghrelin-secreting cells as food-entrainable circadian clocks [published correction appears in *Proc Natl*

Acad Sci U S A 2009;106:17241–17242]. *Proc Natl Acad Sci U S A* 2009; 106: 13582–13587.

20. Cummings DE, Weigle DS, Frayo RS, et al. Plasma ghrelin levels after diet-induced weight loss or gastric bypass surgery. *N Engl J Med* 2002; 346: 1623–1630.
21. Frühbeck G, Diez-Caballero A, Gómez-Ambrosi J, et al. Disruption of the leptin-insulin relationship in obese men 24 hours after laparoscopic adjustable silicone gastric banding. *Obes Surg* 2002; 12: 366–371.
22. Frühbeck G, Diez Caballero A, Gil MJ. Fundus functionality and ghrelin concentrations after bariatric surgery. *N Engl J Med* 2004; 350: 308–309.
23. Oliván B, Teixeira J, Bose M, et al. Effect of weight loss by diet or gastric bypass surgery on peptide YY3-36 levels. *Ann Surg* 2009; 249: 948–953.
24. Bose M, Machineni S, Oliván B, et al. Superior appetite hormone profile after equivalent weight loss by gastric bypass compared to gastric banding. *Obesity (Silver Spring)* 2010; 18: 1085–1091.
25. Korner J, Inabnet W, Febres G, et al. Prospective study of gut hormone and metabolic changes after adjustable gastric banding and Roux-en-Y gastric bypass. *Int J Obes (Lond)* 2009; 33: 786–795.
26. Peterli R, Wölnerhanssen B, Peters T, et al. Improvement in glucose metabolism after bariatric surgery: comparison of laparoscopic Roux-en-Y gastric bypass and laparoscopic sleeve gastrectomy: a prospective randomized trial. *Ann Surg* 2009; 250: 234–241.
27. Langer FB, Reza Hoda MA, Bohdjalian A, et al. Sleeve gastrectomy and gastric banding: effects on plasma ghrelin levels. *Obes Surg* 2005; 15: 1024–1029.
28. Cohen R, Uzzan B, Bihan H, et al. Ghrelin levels and sleeve gastrectomy in super-super-obesity. *Obes Surg* 2005; 15: 1501–1502.
29. Orskov C, Rabenhøj L, Wettergren A, et al. Tissue and plasma concentrations of amidated and glycine-extended glucagon-like peptide I in humans. *Diabetes* 1994; 43: 535–539.
30. Bullock BP, Heller RS, Habener JF. Tissue distribution of messenger ribonucleic acid encoding the rat glucagon-like peptide-1 receptor. *Endocrinology* 1996; 137: 2968–2978.
31. Neary NM, Small CJ, Druce MR, et al. Peptide YY3-36 and glucagon-like peptide-17-36 inhibit food intake additively. *Endocrinology* 2005; 146: 5120–5127.
32. Young AA, Gedulin BR, Rink TJ. Dose-responses for the slowing of gastric emptying in a rodent model by glucagon-like peptide (7-36) NH2, amylin, cholecystokinin, and other possible regulators of nutrient uptake. *Metabolism* 1996; 45: 1–3.
33. Kreymann B, Williams G, Ghatei MA, Bloom SR. Glucagon-like peptide-1 7-36: a physiological incretin in man. *Lancet* 1987; 2: 1300–1304.
34. Soltani N, Kumar M, Glinka Y, et al. In vivo expression of GLP-1/IgG-Fc fusion protein enhances beta-cell mass and protects against streptozotocin-induced diabetes. *Gene Ther* 2007; 14: 981–988.
35. Nauck MA, Homberger E, Siegel EG, et al. Incretin effects of increasing glucose loads in man calculated from venous insulin and C-peptide responses. *J Clin Endocrinol Metab* 1986; 63: 492–498.
36. Le Roux CW, Welbourn R, Werling M, et al. Gut hormones as mediators of appetite and weight loss after Roux-en-Y gastric bypass. *Ann Surg* 2007; 246: 780–785.
37. Morínigo R, Lacy AM, Casamitjana R, et al. GLP-1 and changes in glucose tolerance following gastric bypass surgery in morbidly obese subjects. *Obes Surg* 2006; 16: 1594–1601.
38. Dixon JB, O'Brien PE, Playfair J, et al. Adjustable gastric banding and conventional therapy for type 2 diabetes: a randomized controlled trial. *JAMA* 2008; 299: 316–323.
39. Korner J, Inabnet W, Conwell IM, et al. Differential effects of gastric bypass and banding on circulating gut hormone and leptin levels. *Obesity (Silver Spring)* 2006; 14: 1553–1561.
40. Adrian TE, Ferri GL, Bacarese-Hamilton AJ, et al. Human distribution and release of a putative new gut hormone, peptide YY. *Gastroenterology* 1985; 89: 1070–1077.
41. Batterham RL, Cowley MA, Small CJ, et al. Gut hormone PYY(3-36) physiologically inhibits food intake. *Nature* 2002; 418: 650–654.
42. Batterham RL, Cohen MA, Ellis SM, et al. Inhibition of food intake in obese subjects by peptide YY3-36. *N Engl J Med* 2003; 349: 941–948.
43. Koda S, Date Y, Murakami N, et al. The role of the vagal nerve in peripheral PYY3-36-induced feeding reduction in rats. *Endocrinology* 2005; 146: 2369–2375.
44. Le Roux CW, Aylwin SJ, Batterham RL, et al. Gut hormone profiles following bariatric surgery favor an anorectic state, facilitate weight loss, and improve metabolic parameters. *Ann Surg* 2006; 243: 108–114.
45. Karamanakos SN, Vagenas K, Kalfarentzos F, Alexandrides TK. Weight loss, appetite suppression, and changes in fasting and postprandial ghrelin and peptide-YY levels after Roux-en-Y gastric bypass and sleeve gastrectomy: a prospective, double blind study. *Ann Surg* 2008; 247: 401–407.
46. Valderas JP, Irribarra V, Boza C, et al. Medical and surgical treatments for obesity have opposite effects on peptide YY and appetite: a prospective study controlled for weight loss. *J Clin Endocrinol Metab* 2010; 95: 1069–1075.
47. Polonsky KS, Given BD, Hirsch L, et al. Quantitative study of insulin secretion and clearance in normal and obese subjects. *J Clin Invest* 1988; 81: 435–441.
48. Lee WJ, Ser KH, Chong K, et al. Laparoscopic sleeve gastrectomy for diabetes treatment in nonmorbidly obese patients: efficacy and change of insulin secretion. *Surgery* 2010; 147: 664–669.
49. Zhang Y, Proenca R, Maffel M, et al. Positional cloning of the mouse obese gene and its human homologue [published correction appears in *Nature* 1995;374:479]. *Nature* 1994; 372: 425–432.
50. English PJ, Coughlin SR, Hayden K, et al. Plasma adiponectin increases postprandially in obese, but not in lean, subjects. *Obes Res* 2003; 11: 839–844.
51. Kubota N, Terauchi Y, Yamauchi T, et al. Disruption of adiponectin causes insulin resistance and neointimal formation. *J Biol Chem* 2002; 277: 25863–25866.
52. Weyer C, Funahashi T, Tanaka S, et al. Hypoadiponectinemia in obesity and type 2 diabetes: close association with insulin resistance and hyperinsulinemia. *J Clin Endocrinol Metab* 2001; 86: 1930–1935.
53. Spranger J, Kroke A, Möhlig M, et al. Adiponectin and protection against type 2 diabetes mellitus [published

correction appears in *Lancet* 2003;361:1060]. *Lancet* 2003; 361: 226–228.

54. Kotidis EV, Koliakos G, Papavramidis TS, Papavramidis ST. The effect of biliopancreatic diversion with pylorus-preserving sleeve gastrectomy and duodenal switch on fasting serum ghrelin, leptin and adiponectin levels: is there a hormonal contribution to the weight-reducing effect of this procedure? *Obes Surg* 2006; 16: 554–559.
55. De Paula AL, Stival AR, Macedo A, et al. Prospective randomized controlled trial comparing 2 versions of laparoscopic ileal interposition associated with sleeve gastrectomy for patients with type 2 diabetes with BMI 21-34 kg/m(2). *Surg Obes Relat Dis* 2010; 6: 296–304.
56. Ferzli GS, Dominique E, Ciaglia M, et al. Clinical improvement after duodenojejunal bypass for nonobese type 2 diabetes despite minimal improvement in glycemic homeostasis. *World J Surg* 2009; 33: 972–979.
57. Snyder B, Nguyen A, Scarbourough T, et al. Comparison of those who succeed in losing significant excessive weight after bariatric surgery and those who fail. *Surg Endosc* 2009; 23: 2302–2306.
58. Lin E, Gletsu N, Fugate K, et al. The effects of gastric surgery on systemic ghrelin levels in the morbidly obese. *Arch Surg* 2004; 139: 780–784.
59. Couce ME, Cottam D, Esplen J, et al. Is ghrelin the culprit for weight loss after gastric bypass surgery? A negative answer. *Obes Surg* 2006; 16: 870–878.
60. Sundbom M, Holdstock C, Engström BE, Karlsson FA. Early changes in ghrelin following Roux-en-Y gastric bypass: influence of vagal nerve functionality? *Obes Surg* 2007; 17: 304–310.
61. Borg CM, le Roux CW, Ghatei MA, et al. Progressive rise in gut hormone levels after Roux-en-Y gastric bypass suggests gut adaptation and explains altered satiety. *Br J Surg* 2006; 93: 210–215.
62. Morínigo R, Casamitjana R, Moizé V, et al. Short-term effects of gastric bypass surgery on circulating ghrelin levels. *Obes Res* 2004; 12: 1108–1116.
63. Morínigo R, Vidal J, Lacy AM, et al. Circulating peptide YY, weight loss, and glucose homeostasis after gastric bypass surgery in morbidly obese subjects. *Ann Surg* 2008; 247: 270–275.
64. Frühbeck G, Rotellar F, Hernández-Lizoain JL, et al. Fasting plasma ghrelin concentrations 6 months after gastric bypass are not determined by weight loss or changes in insulinemia. *Obes Surg* 2004; 14: 1208–1215.
65. Holdstock C, Engström BE, Ohrvall M, et al. Ghrelin and adipose tissue regulatory peptides: effect of gastric bypass surgery in obese humans. *J Clin Endocrinol Metab* 2003; 88: 3177–3183.
66. Vendrell J, Broch M, Vilarrasa N, et al. Resistin, adiponectin, ghrelin, leptin, and proinflammatory cytokines: relationships in obesity. *Obes Res* 2004; 12: 962–971.
67. Whitson BA, Leslie DB, Kellogg TA, et al. Entero-endocrine changes after gastric bypass in diabetic and nondiabetic patients: a preliminary study. *J Surg Res* 2007; 141: 31–39.
68. Garcia-Fuentes E, Garrido-Sanchez L, Garcia-Almeida JM, et al. Different effect of laparoscopic Roux-en-Y gastric bypass and open biliopancreatic diversion of Scopinaro on serum PYY and ghrelin levels. *Obes Surg* 2008; 18: 1424–1429.
69. García de la Torre N, Rubio MA, Bordiú E, et al. Effects of weight loss after bariatric surgery for morbid obesity on vascular endothelial growth factor-A, adipocytokines, and insulin. *J Clin Endocrinol Metab* 2008; 93: 4276–4281.
70. Pardina E, López-Tejero MD, Llamas R, et al. Ghrelin and apolipoprotein AIV levels show opposite trends to leptin levels during weight loss in morbidly obese patients. *Obes Surg* 2009; 19: 1414–1423.
71. Faraj M, Havel PJ, Phélis S, et al. Plasma acylation-stimulating protein, adiponectin, leptin, and ghrelin before and after weight loss induced by gastric bypass surgery in morbidly obese subjects. *J Clin Endocrinol Metab* 2003; 88: 1594–1602.
72. Roth CL, Reinehr T, Schernthaner GH, et al. Ghrelin and obestatin levels in severely obese women before and after weight loss after Roux-en-Y gastric bypass surgery. *Obes Surg* 2009; 19: 29–35.
73. Stoeckli R, Chanda R, Langer I, Keller U. Changes of body weight and plasma ghrelin levels after gastric banding and gastric bypass. *Obes Res* 2004; 12: 346–350.
74. Carroll JF, Franks SF, Smith AB, Phelps DR. Visceral adipose tissue loss and insulin resistance 6 months after laparoscopic gastric banding surgery: a preliminary study. *Obes Surg* 2009; 19: 47–55.
75. Hanusch-Enserer U, Brabant G, Roden M. Ghrelin concentrations in morbidly obese patients after adjustable gastric banding. *N Engl J Med* 2003; 348: 2159–2160.
76. Hanusch-Enserer U, Cauza E, Brabant G, et al. Plasma ghrelin in obesity before and after weight loss after laparoscopical adjustable gastric banding. *J Clin Endocrinol Metab* 2004; 89: 3352–3358.
77. Schindler K, Prager G, Ballaban T, et al. Impact of laparoscopic adjustable gastric banding on plasma ghrelin, eating behaviour and body weight. *Eur J Clin Invest* 2004; 34: 549–554.
78. Shak JR, Roper J, Perez-Perez GI, et al. The effect of laparoscopic gastric banding surgery on plasma levels of appetite-control, insulinotropic, and digestive hormones [published correction appears in *Obes Surg* 2008;18:1097–1098]. *Obes Surg* 2008; 18: 1089–1096.
79. Wang Y, Liu J. Plasma ghrelin modulation in gastric band operation and sleeve gastrectomy. *Obes Surg* 2009; 19: 357–362.
80. Adami GF, Cordera R, Marinari G, et al. Plasma ghrelin concentration in the short-term following biliopancreatic diversion. *Obes Surg* 2003; 13: 889–892.
81. Mingrone G, Granato L, Valera-Mora E, et al. Ultradian ghrelin pulsatility is disrupted in morbidly obese subjects after weight loss induced by malabsorptive bariatric surgery. *Am J Clin Nutr* 2006; 83: 1017–1024.
82. Mingrone G, Nolfe G, Gissey GC, et al. Circadian rhythms of GIP and GLP1 in glucose-tolerant and in type 2 diabetic patients after biliopancreatic diversion. *Diabetologia* 2009; 52: 873–881.
83. García-Unzueta MT, Fernández-Santiago R, Domínguez-Díez A, et al. Fasting plasma ghrelin levels increase progressively after biliopancreatic diversion: one-year follow-up. *Obes Surg* 2005; 15: 187–190.

84. Adami GF, Cordera R, Andraghetti G, et al. Changes in serum ghrelin concentration following biliopancreatic diversion for obesity. *Obes Res* 2004; 12: 684–687.
85. Valera Mora ME, Manco M, Capristo E, et al. Growth hormone and ghrelin secretion in severely obese women before and after bariatric surgery. *Obesity (Silver Spring)* 2007; 15: 2012–2018.
86. Bohdjalian A, Langer FB, Shakeri-Leidenmühler S, et al. Sleeve gastrectomy as sole and definitive bariatric procedure: 5-year results for weight loss and ghrelin. *Obes Surg* 2010; 20: 535–540.
87. Kashyap SR, Daud S, Kelly KR, et al. Acute effects of gastric bypass versus gastric restrictive surgery on beta-cell function and insulinotropic hormones in severely obese patients with type 2 diabetes. *Int J Obes (Lond)* 2010; 34: 462–471.
88. Clements RH, Gonzalez QH, Long CI, et al. Hormonal changes after Roux-en-Y gastric bypass for morbid obesity and the control of type-II diabetes mellitus. *Am Surg* 2004; 70: 1–5.
89. Rubino F, Gagner M, Gentileschi P, et al. The early effect of the Roux-en-Y gastric bypass on hormones involved in body weight regulation and glucose metabolism. *Ann Surg* 2004; 240: 236–242.
90. Laferrère B, Heshka S, Wang K, et al. Incretin levels and effect are markedly enhanced 1 month after Roux-en-Y gastric bypass surgery in obese patients with type 2 diabetes. *Diabetes Care* 2007; 30: 1709–1716.
91. Pournaras DJ, Osborne A, Hawkins SC, et al. The gut hormone response following Roux-en-Y gastric bypass: cross-sectional and prospective study. *Obes Surg* 2010; 20: 56–60.
92. Reinehr T, Roth CL, Schernthaner GH, et al. Peptide YY and glucagon-like peptide-1 in morbidly obese patients before and after surgically induced weight loss. *Obes Surg* 2007; 17: 1571–1577.
93. De Carvalho CP, Marin DM, de Souza AL, et al. GLP-1 and adiponectin: effect of weight loss after dietary restriction and gastric bypass in morbidly obese patients with normal and abnormal glucose metabolism. *Obes Surg* 2009; 19: 313–320.
94. Guidone C, Manco M, Valera-Mora E, et al. Mechanisms of recovery from type 2 diabetes after malabsorptive bariatric surgery. *Diabetes* 2006; 55: 2025–2031.
95. Marfella R, Barbieri M, Ruggiero R, et al. Bariatric surgery reduces oxidative stress by blunting 24-h acute glucose fluctuations in type 2 diabetic obese patients. *Diabetes Care* 2010; 33: 287–289.
96. Salinari S, Bertuzzi A, Asnaghi S, et al. First-phase insulin secretion restoration and differential response to glucose load depending on the route of administration in type 2 diabetic subjects after bariatric surgery. *Diabetes Care* 2009; 32: 375–380.
97. Valverde I, Puente J, Martín-Duce A, et al. Changes in glucagon-like peptide-1 (GLP-1) secretion after biliopancreatic diversion or vertical banded gastroplasty in obese subjects. *Obes Surg* 2005; 15: 387–397.
98. Lugari R, Dei Cas A, Ugolotti D, et al. Glucagon-like peptide 1 (GLP-1) secretion and plasma dipeptidyl peptidase IV (DPP-IV) activity in morbidly obese patients undergoing biliopancreatic diversion. *Horm Metab Res* 2004; 36: 111–115.
99. Molina A, Vendrell J, Gutiérrez C, et al. Insulin resistance, leptin and TNF-alpha system in morbidly obese women after gastric bypass. *Obes Surg* 2003; 13: 615–621.
100. Ramos AP, de Abreu MR, Vendramini RC, et al. Decrease in circulating glucose, insulin and leptin levels and improvement in insulin resistance at 1 and 3 months after gastric bypass. *Obes Surg* 2006; 16: 1359–1364.
101. Swarbrick MM, Stanhope KL, Austrheim-Smith IT, et al. Longitudinal changes in pancreatic and adipocyte hormones following Roux-en-Y gastric bypass surgery. *Diabetologia* 2008; 51: 1901–1911.
102. Trakhtenbroit MA, Leichman JG, Algahim MF, et al. Body weight, insulin resistance, and serum adipokine levels 2 years after 2 types of bariatric surgery. *Am J Med* 2009; 122: 435–442.
103. Coughlin CC, Finck BN, Eagon JC, et al. Effect of marked weight loss on adiponectin gene expression and plasma concentrations. *Obesity (Silver Spring)* 2007; 15: 640–645.
104. Riedl M, Vila G, Maier C, et al. Plasma osteopontin increases after bariatric surgery and correlates with markers of bone turnover but not with insulin resistance. *J Clin Endocrinol Metab* 2008; 93: 2307–2312.
105. Serra A, Granada ML, Romero R, et al. The effect of bariatric surgery on adipocytokines, renal parameters and other cardiovascular risk factors in severe and very severe obesity: 1-year follow-up. *Clin Nutr* 2006; 25: 400–408.
106. Swarbrick MM, Austrheim-Smith IT, Stanhope KL, et al. Circulating concentrations of high-molecular-weight adiponectin are increased following Roux-en-Y gastric bypass surgery. *Diabetologia* 2006; 49: 2552–2558.
107. Vila G, Riedl M, Maier C, et al. Plasma MR-proADM correlates to BMI and decreases in relation to leptin after gastric bypass surgery. *Obesity (Silver Spring)* 2009; 17: 1184–1188.
108. Vilarrasa N, Vendrell J, Sánchez-Santos R, et al. Effect of weight loss induced by gastric bypass on proinflammatory interleukin-18, soluble tumour necrosis factor-alpha receptors, C-reactive protein and adiponectin in morbidly obese patients. *Clin Endocrinol (Oxf)* 2007; 67: 679–686.
109. Uzun H, Zengin K, Taskin M, et al. Changes in leptin, plasminogen activator factor and oxidative stress in morbidly obese patients following open and laparoscopic Swedish adjustable gastric banding. *Obes Surg* 2004; 14: 659–665.
110. Moschen AR, Molnar C, Wolf AM, et al. Effects of weight loss induced by bariatric surgery on hepatic adipocytokine expression. *J Hepatol* 2009; 51: 765–777.
111. Ram E, Vishne T, Maayan R, et al. The relationship between BMI, plasma leptin, insulin and proinsulin before and after laparoscopic adjustable gastric banding. *Obes Surg* 2005; 15: 1456–1462.
112. Diker D, Vishne T, Maayan R, et al. Impact of gastric banding on plasma adiponectin levels. *Obes Surg* 2006; 16: 1057–1061.
113. Haider DG, Schindler K, Schaller G, et al. Increased plasma visfatin concentrations in morbidly obese subjects are reduced after gastric banding. *J Clin Endocrinol Metab* 2006; 91: 1578–1581.

114. Engl J, Bobbert T, Ciardi C, et al. Effects of pronounced weight loss on adiponectin oligomer composition and metabolic parameters. *Obesity (Silver Spring)* 2007; 15: 1172–1178.
115. Poitou C, Lacorte JM, Coupaye M, et al. Relationship between single nucleotide polymorphisms in leptin, IL6 and adiponectin genes and their circulating product in morbidly obese subjects before and after gastric banding surgery. *Obes Surg* 2005; 15: 11–23.
116. Salinari S, Bertuzzi A, Iaconelli A, et al. Twenty-four hour insulin secretion and beta cell NEFA oxidation in type 2 diabetic, morbidly obese patients before and after bariatric surgery. *Diabetologia* 2008; 51: 1276–1284.
117. Camastra S, Manco M, Frascerra S, et al. Daylong pituitary hormones in morbid obesity: effects of bariatric surgery. *Int J Obes (Lond)* 2009; 33: 166–172.
118. Rosa G, Mingrone G, Manco M, et al. Molecular mechanisms of diabetes reversibility after bariatric surgery. *Int J Obes (Lond)* 2007; 31: 1429–1436.
119. Bruno C, Fulford AD, Potts JR, et al. Serum markers of bone turnover are increased at six and 18 months after Roux-en-Y bariatric surgery: correlation with the reduction in leptin. *J Clin Endocrinol Metab* 2010; 95: 159–166.
120. Das SK, Roberts SB, McCrory MA, et al. Long-term changes in energy expenditure and body composition after massive weight loss induced by gastric bypass surgery. *Am J Clin Nutr* 2003; 78: 22–30.
121. Czupryniak L, Pawlowski M, Kumor A, et al. Predicting maximum Roux-en-Y gastric bypass-induced weight reduction–preoperative plasma leptin or body weight? *Obes Surg* 2007; 17: 162–167.
122. Maruna P, Gürlich R, Fried M, et al. Leptin as an acute phase reactant after non-adjustable laparoscopic gastric banding. *Obes Surg* 2001; 11: 609–614.
123. Heinonen MV, Purhonen AK, Miettinen P, et al. Apelin, orexin-A and leptin plasma levels in morbid obesity and effect of gastric banding. *Regul Pept* 2005; 130: 7–13.

Office-based Management of Obesity

Ken Fujioka, MD and Nitasha Bakhru, MD

Scripps Clinic Division of Diabetes and Endocrinology, San Diego, CA

OUTLINE

How Can We Treat the Disease If It Is Not Diagnosed?
Examine Not Only Your Patients, But Your Approach to Them
The Metabolic Hurdle: Why Is It So Hard to Lose Weight?
Show Me The Money

ABSTRACT

Appropriate recognition and management of obesity is imperative given the fact that one-third of Americans are clinically obese. Up to 25% of patients with a body mass index ≥30 are not appropriately identified by primary care physicians as being obese. An even smaller portion of recognized patients is offered any form of weight-loss treatment. This may be the result of approximately 40% of physicians possessing negative attitudes toward such patients given the challenges in managing this population. In the near future, physicians will be empowered with further knowledge regarding metabolic and neurologic adaptations that make weight loss so difficult. Along with this new information will come novel treatment options. It is predicted that many of the barriers preventing appropriate management of obesity will diminish as further knowledge and treatment options become available. ***Mt Sinai J Med 77:466–471, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: obesity, office-based management, under diagnosis, weight-loss therapy.

Address Correspondence to:

Ken Fujioka
Scripps Clinic Division of
Diabetes and Endocrinology
San Diego, CA

Email: fujioka.ken@scrippshealth.org

Obesity is a worldwide public health concern. In the United States alone, more than 133 million people, or 68% of the population, can be categorized as overweight or obese as defined by a body mass index (BMI) ≥25. Data obtained from the 2007–2008 National Health and Nutrition Examination Survey (NHANES) revealed that 33.8% of Americans can be further classified as obese (BMI ≥30), with women constituting a slightly larger fraction of the total studied.[1]

These data are astounding when compared with data obtained between 1988 and 1994, when the prevalence of obesity was only 23%. Furthermore, the prevalence of obesity on a global scale continues to rise, but it is difficult to quantify given differing BMI classification schemes in Asian populations. Obesity has become a health concern not only in industrialized nations but also in underserved countries, where it is starting to rival malnutrition, resulting in comorbidities of similar magnitude.[2] As the number of obese individuals has increased, so has the severity of the disease. Obesity-related illnesses such as diabetes are on the increase, with the incidence of diabetes mellitus (DM) escalating from 5.0% in 1991 to 8.1% in 2002. It is estimated that, in the United States, the lifetime risk of developing DM is 1 in 3 persons (Figures 1 and 2).[3]

> *It is estimated that, in the United States, the lifetime risk of developing diabetes mellitus is 1 in 3 persons.*

Even more concerning is that such findings have been noted in the pediatric population as well. NHANES data obtained between 2003 and 2006, compared with data from 1976 to 1980, revealed that the prevalence of obesity (BMI ≥95th percentile) increased from 5.0% to 17.6% in children aged 12–19 years.[4] Even more concerning is data showing that the rate of death among children in the highest BMI quartile is twice that of their cohorts in the lowest BMI quartile. Furthermore, the rate of

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20201

© 2010 Mount Sinai School of Medicine

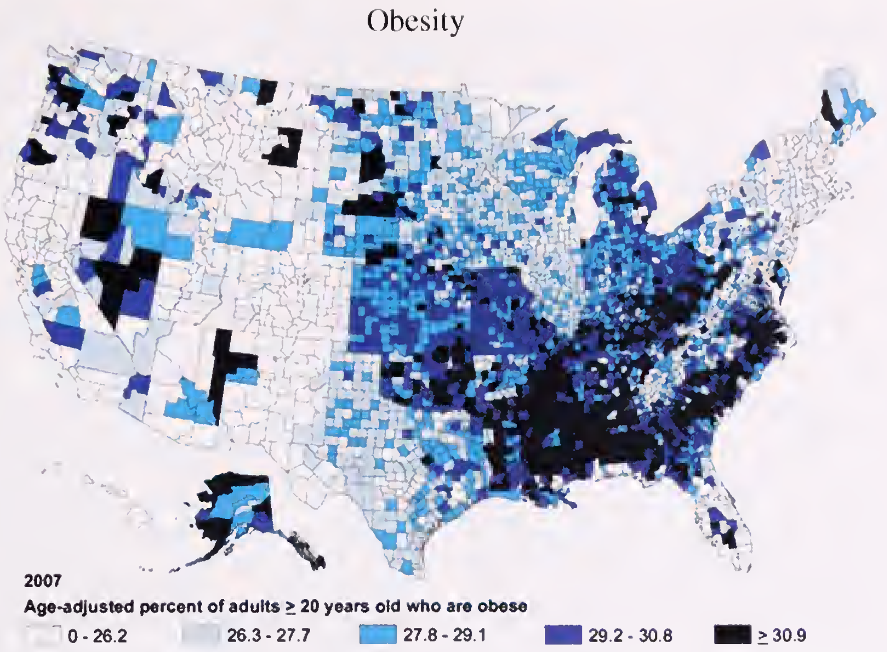

Obesity
2007
Age-adjusted percent of adults ≥ 20 years old who are obese
0 - 26.2
26.3 - 27.7
27.8 - 29.1
29.2 - 30.8
≥ 30.9

Fig 1. Estimated county-level prevalence of obesity–United States, 2007. Morbidity and Mortality Weekly Report, November 20, 2009;58:1259–1263. http://www.cdc.gov/obesity/data/trends.html.

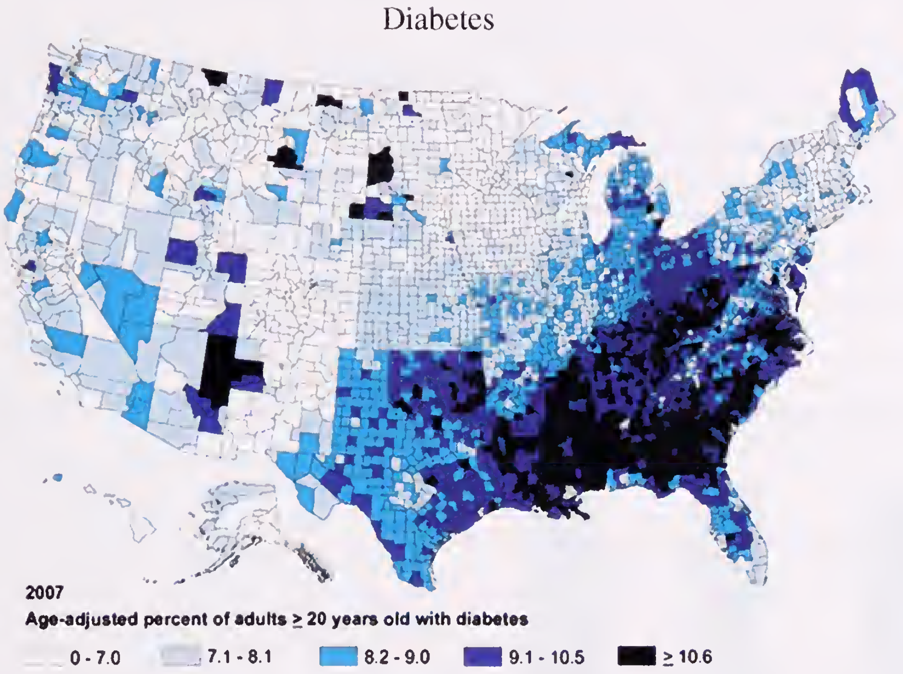

Diabetes
2007
Age-adjusted percent of adults ≥ 20 years old with diabetes
0 - 7.0
7.1 - 8.1
8.2 - 9.0
9.1 - 10.5
≥ 10.6

Fig 2. Estimated county-level prevalence of diabetes–United States, 2007. Morbidity and Mortality Weekly Report, November 20, 2009;58:1259–1263. http://www.cdc.gov/obesity/data/trends.html.

death in children in the highest quartile of glucose intolerance is 73% higher than that of children in the lowest quartile. Such studies emphasize the association between childhood obesity and premature death.[5]

HOW CAN WE TREAT THE DISEASE IF IT IS NOT DIAGNOSED?

What can we do to combat obesity? Perhaps the first step is recognition of the disease, followed by knowledge of potential treatment modalities. The Primary Care Obesity Diagnosis Project sought to evaluate the frequency of appropriate recognition of obesity among primary care physicians at the Mayo Clinic in Rochester, MN. Evaluation of the primary care database revealed that of the 9827 patients seen for an annual exam between November 2004 and October 2005, 2543 were categorized as obese based upon a calculated BMI $\geq$30. Surprisingly, of those who qualified as being obese, only 505 (19.9%) had a diagnosis of obesity documented somewhere in their chart (based on an extensive chart review, not ICD-9 coding alone). Obese elderly patients and men were significantly less likely to be properly identified and diagnosed. However, those with a BMI $\geq$35 or with comorbidities such as DM and obstructive sleep apnea were more likely to be diagnosed. The unifying theme underlying the study was that the diagnosis of obesity was the strongest predictor of formulation of an obesity management plan (Figure 3).[6]

Recognition of overweight and obesity needs to be bolstered in multiple parts of the country. Identification of overweight status requires improvement among not only patients, but also physicians. A study comparing physician and patient perception of weight status and BMI was conducted among 788 patients in an internal medicine community-based practice. Appropriately, 80% of overweight patients correctly perceived their weight status, and 75% of such patients were accurately identified by their physicians. This suggests that although both parties are accurately identifying weight status the majority of the time, there is still room for improvement. We cannot treat the disease if it is not diagnosed (Table 1).[7]

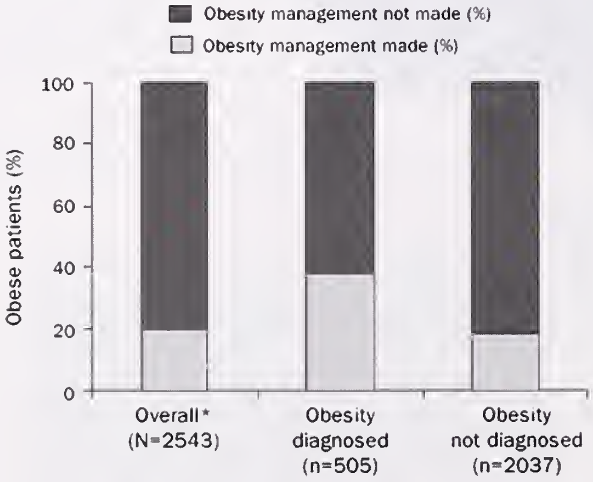


Fig 3. Documentation of obesity management among obese patients and impact of obesity diagnosis on management. *One patient had missing data.

Table 1. *Body Mass Index of Overweight Patients Who Were Misclassified as Normal Weight by the Patients or Their Physicians.*

Body Mass Index (kg/m^2)	Patient Errors (n = 108)	Physician Errors (n = 132)
	Number (%)	
25.0–26.9	42 (39)	71 (54)
27.0–29.9	42 (39)	42 (32)
$\geq$30 (obese)	24 (22)	19 (14)

We cannot treat the disease if it is not diagnosed. The Primary Care Obesity Diagnosis Project, a 2004–2005 study by the Mayo Clinic involving 9827 patients, showed that a diagnosis of obesity was the strongest predictor that an obesity management plan would be formulated.

The array of treatment modalities for the obese patient has provided both options and also frustration for the primary care physician. A level of discomfort with the numerous modalities may be a hindrance to their adoption and application. Such options range from lifestyle intervention to medical therapy and bariatric surgery. Patients with a BMI $\geq$30, or $\geq$27 with a weight-related health problem, qualify for medical pharmacotherapy. The use of pharmacotherapy in conjunction with diet, exercise, and behavior modification is more efficacious than traditional dieting alone, which typically results in <5% weight loss long-term. Interventions completed under the guidance of a physician utilizing a multispecialty approach (involving a dietitian, possibly a bariatric surgeon, and a psychologist) are thought to be more long-lasting than those pursued through commercial nonmedical weight-loss centers.[8] Bariatric surgery, indicated for individuals with BMI >40 (or >35 with obesity-related comorbidities), has the highest success rate and can often be completed laparoscopically. Unfortunately, many patients who would benefit from bariatric surgery are not aware of this option. Perhaps it is a lack of recognition of their weight status that is responsible for this.

Many patients who would benefit from bariatric surgery are not aware of this option.

Many methods to increase awareness of obesity among physicians have been attempted. A study conducted among family practitioners investigated the efficacy of a prompt appearing in patients' electronic charts alerting physicians of their BMI. Obese patients of physicians who had a BMI chart prompt were significantly more likely than obese patients who did not receive a prompt in their medical record to receive a diagnosis of obesity. These patients were also more likely to receive referral for diet treatment (14.0% versus 7.3%, $P = 0.002$) and exercise (12.1% versus 7.1%, $P = 0.016$).[9]

Multiple barriers exist hindering recognition of obesity and subsequent formulation of a management plan. Lack of physician recognition is driven by factors such as lack of time available to counsel, lack of knowledge, and frustration from prior attempts. Patients are impaired by factors such as lack of motivation, lack of knowledge, and fear of embarrassment. The medical system, in general, is limited by lack of resources and rising cost. Although many factors exist, the one unifying theme is that lack of knowledge exists among all involved parties.[6] Confronting this barrier may be the first step to success.

If recognition is the first step toward improvement, perhaps integration of BMI as a vital sign is our most powerful tool. BMI has been shown to have an inverse relationship with overall patient self-reported health, especially when it exceeds 35 kg/m^2. A patient-centered decision rule for referral of those with BMI ≥ 35 has been suggested as a means to identify and treat those most at risk.[10] Focused training on the subject for physicians at large may improve acceptance on part of both the practitioner and involved patients. Behavioral therapy to modify patients' perceptions of necessary lifestyle modifications may also prove efficacious, especially when integrated into a model of weight-centric office visits or referral to a weight-management clinic. Multiple modalities of alleviating the impact of obesity on our population exist, but they can only be employed after appropriate recognition of the disease occurs.

Multiple modalities of alleviating the impact of obesity on our population exist, but they can only be employed after appropriate recognition of the disease occurs.

EXAMINE NOT ONLY YOUR PATIENTS, BUT YOUR APPROACH TO THEM

Unfortunately, 30% to 40% of physicians will have a negative attitude and reaction toward the obese patient.[11,12] Even healthcare professionals who specialize in obesity are not immune to negative attitudes.[13] Physicians may be biased by their own personal experiences with weight loss, prior patient encounters, or cultural views. Furthermore, the field of obesity derives its roots from nonobjective sources as well. Prior to the 1995 discovery of leptin, a hormone responsible for regulating body weight, very little was known regarding the physiology of obesity. As a result, many of the initial treatment options had poor science behind them and were characterized by undesirable side effects.[14] Thus, it is not surprising that many physicians became biased toward not only the obese patient, but also potential treatment options.

Fortunately, there has been an explosion of scientific knowledge and discovery regarding the science of weight regulation in humans.[15] This has resulted in a heightened understanding of the physiological challenges involved in the achievement of weight loss. Consequently, more treatment options have become available that provide innovative, targeted approaches to potentiating weight loss. The array of pharmacotherapeutic and surgical options in particular has grown.

Prior studies have also examined why physicians are biased toward treating the obese patient. What has become apparent is lack of knowledge and comfort with the field.[12] As more advances are made within the field of weight management, perhaps this bias will dissipate as physicians are empowered by an expanding legion of evidence-based treatment algorithms.

THE METABOLIC HURDLE: WHY IS IT SO HARD TO LOSE WEIGHT?

With the discovery of the feedback loops regulating weight status, we now have an expanded knowledge base regarding weight regulation. We now understand that even after as little as a 5% to 10% loss of body weight, numerous compensatory mechanisms are activated to halt the process of further weight loss.[15] Neural pathways are turned on to stimulate eating and lower the ability to burn calories. Patients will commonly refer to this as a "plateau." Although patients continue to diet, their weight loss slows down or even stops, despite their best efforts.

This fight by the human body to defend against weight loss is further hampered by our environment. There is little doubt we live in a society of abundant food. Combining our modern environment with the human body's conventional physiology to defend against weight loss often results in a net weight gain. With this knowledge in mind, healthcare providers need to be very careful when discussing weight-loss goals. A goal of 10% weight loss is reasonable and within many patients' capabilities. Such a modest goal has been shown to be beneficial in reducing the impact of metabolic derangements occurring as a result of obesity. Improvements in glycemic control, dyslipidemia, and hypertension can be noted with such a range of weight loss.[16] However, many patients will have the initial desire to lose much more than 10% despite our best efforts at education. In a study examining patient expectations, patients were noted to have the desire to lose 25% of their weight despite what their healthcare provider suggested.[17]

SHOW ME THE MONEY

Oftentimes the most challenging aspect of practicing obesity medicine is obtaining insurance coverage for treatment options. Both healthcare providers and patients become disgruntled when a claim for a visit is denied because "obesity treatment" is not covered. However, there are situations where insurance coverage can be successfully attained in many obese patients.

To find these situations, one should treat obesity like any other medical condition and define the comorbid conditions. A good example is Diabetes Mellitus type 2 (DM) (ICD-9 code 250), or even impaired fasting glucose (fasting blood sugars 100–125 mg/dL, ICD-9 code 790.1). The treatment for DM is to control blood sugar, and it has been shown that by losing weight the blood sugar will improve. Thus, coding the visit for DM is very appropriate. In most instances, the dietitian will also be covered. This frees up the physician by referring the patient to healthcare personnel who can spend time discussing food intake.

> *Oftentimes the most challenging aspect of practicing obesity medicine is obtaining insurance coverage for treatment options. However, there are situations where insurance coverage can be successfully attained in many obese patients. To find these situations, one should treat obesity like any other medical condition and define the comorbid conditions.*

Unfortunately, if a patient has isolated obesity without any comorbidities it can be very difficult to get coverage for the visit or for weight-loss medications. However, the one caveat to this is if a patient is morbidly obese. Morbid obesity is defined as a BMI ≥ 40, or alternatively >35 when a significant comorbidity such as diabetes is present (ICD-9 code 278.01). These patients qualify for obesity management given that weight loss would provide a considerable benefit in their health status. The knowledge of the potential barriers to success, and methods to overcome them, should enable healthcare providers to become more comfortable in treating the ever-expanding population of obese individuals.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999–2008. *JAMA* 2010; 303: 235–241.
2. World Health Organization. Global Strategy on Diet, Physical Activity, and Health. http://www.who.int/nutrition/topics/obesity/en/index.html.
3. Narayan KM, Boyle JP, Thompson TJ, et al. Lifetime risk for diabetes mellitus in the United States. *JAMA* 2003; 290: 1884–1890.
4. Ogden CL, Carroll MD, Flegal KM. High body mass index for age among US children and adolescents, 2003–2006. *JAMA* 2008; 299: 2401–2405.
5. Franks PW, Hanson RL, Knowler RC, et al. Childhood obesity, other cardiovascular risk factors, and premature death. *N Engl J Med* 2010; 362: 485–493.
6. Bardia A, Holtan SG, Slezak JM, Thompson WG. Diagnosis of obesity by primary care physicians and impact on obesity management. *Mayo Clin Proc* 2007; 82: 927–932.
7. Caccamese SM, Kolodner K, Wright SM. Comparing patient and physician perception of weight status with body mass index. *Am J Med* 2002; 112: 662–666.
8. Fujioka K. Weight loss clinic. In: Bray GA, Bouchard C, eds. *Handbook of Obesity: Clinical Applications*. 3rd ed. New York, NY: Informa Healthcare; 2008; 593–606.
9. Schriefer SP, Landis SE, Turbow DJ, Patch SC. Effect of a computerized body mass index prompt on diagnosis and treatment of adult obesity. *Fam Med* 2009; 41: 502–507.

10. Rohrer JE, Merry SP, Lopez-Jimenez F, et al. A patient-centered decision rule for referral of patients to weight-loss programs. *Qual Manag Health Care* 2007; 16: 250–255.
11. Bocquier A, Verger P, Basdevant A, et al. Overweight and obesity: knowledge, attitudes, and practices of general practitioners in France. *Obes Res* 2005; 13: 787–795.
12. Jay M, Kalet A, Ark T, et al. Physicians' attitudes about obesity and their associations with competency and specialty: a cross-sectional study. *BMC Health Serv Res* 2009; 9: 106.
13. Bray GA, Bouchard C. *Handbook of Obesity: Clinical Applications*. 3rd ed. New York, NY: Informa Healthcare; 2008.
14. Greenway FL, Caruso MK. Safety of obesity drugs. *Expert Opin Drug Saf* 2005; 4: 1083–1095.
15. Schwartz MW, Woods SC, Porte D Jr, et al. Central nervous system control of food intake. *Nature* 2000; 404: 661–671.
16. Fujioka K. Benefits of moderate weight loss in patients with type 2 diabetes. *Diabetes Obes Metab* 2010; 12: 186–194.
17. Wadden TA, Womble LG, Sarwer DB, et al. Great expectations: "I'm losing 25% of my weight no matter what you say." *J Consult Clin Psychol* 2003; 71: 1084–1089.
18. Finkelstein EA, Fiebelkorn IC, Wang G. National medical spending attributable to overweight and obesity: how much, and who's paying? *Health Aff (Millwood)* 2003; W3(suppl Web Exclusives): 219–226.

Psychological Treatments for Obesity in Youth and Adults

Rebecca Lo Presti, MA,[1,2] Justine Lai, BA,[1] Tom Hildebrandt, PsyD,[1] and Katharine L. Loeb, PhD[1,2]

[1]Fairleigh Dickinson University, Teaneck, NJ
[2]Mount Sinai School of Medicine, New York, NY

OUTLINE

Overweight and Obesity in Adults
  Behavioral Weight Loss Interventions
  Cognitive and Cognitive Behavioral Weight Loss Interventions
  Further Considerations
Pediatric Overweight and Obesity
  Behavioral Weight Loss Interventions
  Family-Based Treatment
  Current Innovation in the Prevention and Treatment of Adolescent Obesity
Binge Eating Disorder
Advances in Disseminating Interventions
Conclusion

ABSTRACT

Obesity has become a leading global health concern affecting both physical and mental quality of life. This article reviews the psychological interventions currently utilized in confronting the obesity epidemic. It examines obesity in both the pediatric and adult populations, specifically considering empirical evidence of weight loss interventions through randomized control trials. Behavioral weight loss intervention and family-based interventions emerged as 2 of the more frequently utilized and data-supported approaches. Empirically supported psychological treatments for subgroups of the obese population, such as individuals with binge eating disorder, are considered here as well. Limitations of the current interventions are discussed, including issues with generalizability across demographic variables such as age and severity of overweight status. Despite these limitations, overall, psychological interventions play an important role in achieving and maintaining weight loss. ***Mt Sinai J Med 77:472–487, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: behavioral weight loss interventions, binge eating disorder, childhood obesity, family-based interventions, obesity.

Address Correspondence to:

Tom Hildebrandt
Eating and Weight Disorders Program
Mount Sinai School of Medicine
New York, NY
Email: tom.hildebrandt@mssm.edu

Obesity is a major public-health concern, where there is a pressing need to identify both preventive and treatment measures. There is a long history of applying psychological treatments to obese individuals, although the results have been mixed and there is a range of obstacles. Most clearly, disseminating and implementing these treatments for at-risk and affected individuals in the community poses specific challenges. This article will specifically review the evidence supporting psychological interventions for weight loss and maintenance in adults and youth, will consider the difficulties in implementation and treatment, and aims to review the major

> *Disseminating and implementing psychological treatments for individuals at risk or affected by obesity poses specific challenges.*

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20205

© 2010 Mount Sinai School of Medicine

evidence supporting specialist-based psychological treatments for obesity.

OVERWEIGHT AND OBESITY IN ADULTS

Obesity is the second leading preventable cause of disease and death in the United States and has come to be considered both a national and global health epidemic.[1–3] The World Health Organization (2000) defines obesity as a body mass index (BMI) >30 kg/m^2 and overweight as a BMI >25 kg/m^2. Over the last 45 years, the age-adjusted adult obesity prevalence in the United States has increased from 12.8% to 33.8%, and the combined adult overweight and obesity prevalence has risen from 43.4% to 68%.[4,5] Similar growth rates have been observed in England, and countries in Latin America, the Middle East, and the western Pacific have the highest rates of obesity among lower-income and middle-income countries.[2] Given that obesity increases risk for cardiovascular diseases, type 2 diabetes, stroke, gallbladder disease, asthma, sleep disorders, and other health consequences, resulting in associated healthcare costs of more than $117 billion per year, it is critically important to identify efficacious treatments for weight loss (US Department of Health and Human Services, 2004).[6,7] Along with physical health risks, obesity is also associated with negative psychosocial quality of life, particularly among the severely obese, suggesting that psychological factors are important clinically.[8,9]

Behavioral weight loss (BWL) and cognitive behavioral therapy (CBT) are the 2 most commonly used psychological interventions for weight-loss treatments in adults. Dialectical behavior therapy and interpersonal therapy have also been examined for the subpopulation of obese subjects with binge eating disorder (BED).[10,11] These interventions vary in several important ways, including underlying theoretical focus, setting of delivery, and provider training.

Randomized controlled trials examining weight loss through psychological change interventions have found short-term weight loss of up to approximately 10 kg for 6 months after intervention, with gradual weight regain resulting in <5 kg after 2 to 3 years.[12–14] This is perhaps the most reliable finding in clinical science. Within the different categories of psychological treatments there is no comprehensive validated treatment, so that there is great variability between trials, making comparisons difficult.[15] However, we suggest 3 main treatment categories: behavioral change, cognitive and cognitive behavioral, and specialty treatment for BED.

Behavioral Weight Loss Interventions

BWL interventions utilize basic behavioral principles (eg, operant learning, stimulus control, social learning theory) to break patterns of poor dietary choices and sedentary activity.[16] These interventions use contingencies to increase motivation to change and a mixture of positive and negative reinforcement to change dietary choices and physical activity. Interventions generally consist of multiple face-to-face individual or group counseling sessions over 3 to 6 months and target a variety of dietary and activity-based behaviors. All BWL interventions use some form of self-monitoring, which entails the systematic recording of factors important for weight loss, including caloric intake or activity level. This intervention typically provides the foundation for more advanced interventions such as stimulus control, problem solving, relapse prevention, or goal setting.

BWL is the most commonly researched modality in randomized controlled trials with established efficacy compared with no treatment or diet and exercise programs alone.[13,14,17] Dismantling studies that have examined the specific efficacy of different behavioral interventions generally suggests that certain strategies such as self-monitoring, supportive reinforcement, increased exercise, and problem solving are important ingredients for BWL.[18–21]

> *Self-monitoring, supportive reinforcement, increased exercise, and problem solving are important ingredients for behavioral weight loss.*

However, although the length of treatment, and subsequently the average weight loss per study, have substantially multiplied over time,[22,23] the limited follow-up period of most of these studies (few include >12 months post-treatment) makes it difficult to determine the efficacy of these strategies in long-term weight loss. Among these published studies, those with the longer follow-up were more likely to find marginal or no effects.[21,24,25]

The disappointment in these findings led many to move the focus from acute weight loss to long-term weight maintenance.[12,21,26] Attempts to modify the current standardized behavioral treatment to facilitate long-term maintenance have achieved some delay in rate of weight regain postintervention.

The various forms of these modifications have been examined by researchers. For example, 2 extended therapy programs, relapse prevention training (RPT) and problem solving therapy (PST), were compared with standard BWL with no extended contact in 80 obese women.[27] PST, which utilized a therapist to help groups generate their own solutions to barriers for weight loss, was found to be significantly more effective than standard care in achieving clinically significant weight loss of >10% of body weight after 1 year (35% versus 6%), as opposed to minimal effects for RPT (21%). PST was also found to be effective within a stepped care (SC) model, where 20 obese patients were given additional PST sessions if they failed to meet set goals during the weight-maintenance phase, as opposed to a standardized maintenance program (n = 20).[28] On average, participants in the BWL + SC (PST) condition lost significantly more weight as compared with standard BWL (−9.57 ± 5.08 kg versus −6.62 ± 5.81 kg).

Wing and colleagues[29] incorporated a SC approach into 2 weight-maintenance programs, with the only difference in program content being that one intervention was delivered face-to-face and the other group conducted all contact through the Internet. The face-to-face group (n = 105) regained significantly less weight than the control group (2.5 ± 6.7 kg versus 4.9 ± 6.5 kg), whereas the Internet intervention (n = 104) was effective in decreasing the proportion of participants who regained ≥2.3 kg as compared with the control group (54.8% versus 72.4%), but did not have significant differences in the degree of weight regain (4.7 ± 8.6 kg versus 4.9 ± 8.6 kg). There have been mixed findings on what degree of personal contact may be effective; Svetkey *et al.*[30] also noted differences in weight regain between face-to-face and Internet interventions, although telephone contact was found to be just as effective as face-to-face.[31]

Other methods that have been utilized to increase the effectiveness of weight maintenance have included strategies such as varying exercise prescriptions,[32] manipulating social support,[33,34] and exploring models of food provision,[35] all of which have been found to affect weight loss to varying degrees of success.

The Diabetes Prevention Program (DPP) Lifestyle Intervention for patients at risk for diabetes and the Action for Health in Diabetes (Look AHEAD) trial for patients with type 2 diabetes are the 2 largest randomized controlled trials on intentional weight loss and maintenance, incorporating the successful strategies highlighted by the current available literature within a highly structured and individualized program.[36,37] In the Look AHEAD trial, 55.2% of participants (n = 5145) reduced their weight by at least 7% from baseline after 1 year (n = 2496).[38] Of the 1029 participants in the DPP Lifestyle Intervention, 46% met their weight loss goal to reduce weight by a minimum of 7% from baseline, with an average weight loss of 4.1 kg at 3-year follow-up[39]; however, at 6 years, the average weight loss had decreased to 2 kg.[40] The results of these trials are promising for short-term weight loss, but the long-term results are still sobering, as the vast majority of individuals regain weight over time.

Cognitive and Cognitive Behavioral Weight Loss Interventions

Whereas the behavior-based intervention specifically focuses on changing learned behaviors, CBT interventions for weight loss are focused on changing and adapting distorted cognitions associated with overeating or lack of activity.[41,42] Most of the specific behavioral strategies from existing BWL treatments can be found in CBT weight loss treatments. However, these treatments are more focused on identifying cognitions and emotions that facilitate overeating and inactivity, establishing motivational readiness, managing patient expectations for treatment, and cognitive restructuring, particularly around past weight loss attempts.[43]

When compared with no treatment or diet and exercise alone, CBT-based treatments have been found to be more effective in achieving weight loss outcomes,[44–47] and the effect seems to be augmented when CBT is combined with dietary restriction.[48]

> *When compared with no treatment or diet and exercise alone, variations of cognitive behavioral therapy (CBT) have been found to be more effective in achieving weight loss outcomes, and the effect seems to be augmented when CBT is combined with dietary restriction.*

In a comparison of BWL and CBT in 24 obese women, BWL led to greater acute weight loss, but CBT participants achieved greater average net weight losses at 12-month follow-up post-treatment (−10.06 ± 3.42 kg versus −4.29 ± 2.46 kg).[49] Cooper and Fairburn[42] proposed a new CBT specifically tailored for weight loss maintenance. A large-scale study comparing this version of CBT to BWL is near completion in the United Kingdom, attempting to improve upon the modest impact found in other trials of CBT for obesity.[50] However, Rapoport *et al.*[51] found no differences in outcome between a weight loss–tailored CBT and standard CBT.

It seems that further research on CBT is worth exploring, but the increased cost of administering CBT compared with BWL limits enthusiasm. While BWL can generally be delivered by practitioners with varying degrees of training, correct administration of CBT generally requires training by a specialized program and a graduate degree. Thus, this limits its applicability among community or other settings where resources may not be as readily available. Self-help CBT, either with or without the aid of a therapist or supportive facilitator, has shown promise for the treatment of BED[52–55] and may therefore be a more viable option for long-term dissemination in weight loss treatments.

Further Considerations

There is evidence that successful long-term weight loss is possible. McGuire *et al.*[56] found in a nationally representative telephone survey that 21% of overweight adults who had intentionally lost at least 10% of their body weight at one point had been able to maintain the loss for at least 1 year. Only 55% of registrants in the National Weight Control Registry, where eligible individuals must have maintained a weight loss of ≥30 lb. for ≥1 year, used a formal weight loss program; the remainder lost weight on their own. Wing and Hill[57] examined the behavioral strategies used by successful weight loss maintainers in the National Weight Control Registry and found that frequent self-monitoring of weight, high levels of physical activity, and eating a low-fat diet were the most commonly used techniques. It is unclear what percentage of those who lost weight in a formal weight loss program specifically enrolled in a psychological intervention, as opposed to a commercial weight loss program; Brownell *et al.*[58] found that only 5% of those who reported losing and maintaining a significant amount of weight had used a hospital-based or university-based program, as opposed to commercial programs (20%), diet pills (3%), or self (72%). These findings are more promising than those reported by the randomized controlled trials described above, which may reflect different weight loss measurement methods or the fact that those who seek help for weight loss in clinical settings are not reflective of the general population. The fact that a significant number of those who successfully lose weight are able to achieve this goal without any direct intervention suggests that the key component may lie within the individual rather than be delivered through a clinical intervention.

PEDIATRIC OVERWEIGHT AND OBESITY

Obesity has also become one of the most prevalent health concerns for children. Though the growing media attention surrounding this issue may suggest improvements, childhood obesity continues to increase at an alarming rate. Data collected from the Centers for Disease Control and Prevention show a 4-fold increase in overweight children over the last 3 decades and a 3-fold increase in overweight adolescents.[59] Further, it is estimated that 11.3% of children and adolescents aged 2–19 years are at or above the 97th percentile of BMI for age, 16.3% are at or above the 95th percentile, and 31.9% are at or above the 85th percentile.[60] This is especially alarming given the extensive consequences of pediatric overweight and obesity, including cardiovascular risk factors, adult obesity, and obesity-related comorbidities.[61] It is estimated that approximately 90% of overweight children will go on to become obese adults.[62] In addition, health complications such as elevated cholesterol, hypertension, type 2 diabetes, and sleep apnea, as well as psychosocial problems such as depression, disturbed body image, social isolation, and low self-concept, are associated with continued obesity.[62–67] Overweight status in a pediatric population is conventionally determined with BMI, applying age-specific and sex-specific parameters. Pediatric obesity is commonly defined as a BMI ≥95th percentile for age and sex, whereas children with a BMI within the 85th to 95th percentile are considered overweight.

> *It is estimated that approximately 90% of overweight children will go on to become obese adults.*

In light of the significant scope and consequences of pediatric overweight and obesity, efforts to determine optimal points of intervention and the most efficacious forms of treatment are high public-health priorities. Research suggests that the notion of early intervention in obesity cannot be applied too soon; even the neonatal environment may yield risks for childhood obesity and in turn may be a logical entry point for prevention. For example, Dabelea *et al.*[68] found that children of women who had type 2 diabetes during pregnancy were more obese and more likely to develop type 2 diabetes themselves than their siblings born before the mother developed diabetes. This is of particular concern, as these findings suggest a cyclic nature in the development of obesity: a young woman born to a diabetic mother is likely to develop a similar condition that consequently will be transferred to her children. Findings like this amplify the significant need for immediate and effective preventive and early interventions. But although prevention is optimal, it is also aspirational at this point, and extant obesity in childhood and adolescence remains a significant public-health concern. It is therefore imperative that research focus on effective interventions aimed at treating this continually growing epidemic.

Existing research on interventions for child and adolescent obesity has spanned behavioral/psychological, pharmacological, and surgical options.[69–71] The Cochrane Review[72] has attempted to assess the efficacy of such lifestyle, drug, and surgical interventions for pediatric obesity through review of randomized controlled trials from 1985 to 2008. Though their full review is beyond the scope of this article, their findings are of note. In considering lifestyle interventions for both children and adolescents, the meta-analysis suggests a reduction in overweight at 6-month and 12-month follow-ups, regardless of additional pharmacological interventions.

Within psychological interventions, the main modalities studied in randomized controlled trials are individual behavioral treatment and family-based interventions. The majority of these approaches are focused on weight loss; however, a reasonable goal for younger patients is to arrest a weight-gain trajectory while allowing for height to increase, thereby reducing BMI for age percentile over time. Although psychological treatment approaches vary, 3 major components in most intervention programs for pediatric obesity are diet, exercise, and behavioral techniques, whereby specific behaviors are defined and identified as targets of change.

Behavioral Weight Loss Interventions

BWL interventions specifically target weight loss in the child or adolescent through the use of behavior-modification techniques, which identify maladaptive behaviors and match strategies to effect change. Typical modifications in pediatric weight loss are focused on diet and activity levels. This process may involve measuring the frequency and intensity of the identified behavior, as achieved through self-monitoring, in which the child and/or parent records the behaviors. Following these observations, a behavior-change strategy is developed. This may include stimulus-control techniques based in classical conditioning paradigms, such as limiting exposure to unhealthy food or disassociating eating with other behaviors such as television viewing. Operant conditioning techniques such as contingencies and reinforcement may also be applied; for example, when a targeted behavior is performed, it is positively reinforced through a reward. Additional strategies such as continued self-monitoring, environmental management (ie, eating in a designated place), and modeling (ie, having parents behave as the child is expected to behave) are also implemented to create change.[73]

The short-term efficacy of behavior modification for pediatric overweight and obese populations has been documented in groups of children aged ≤12 years with decreases in percent overweight of 5% to 20%. Interventions targeting weight-control behaviors such as modification of eating and physical activity have demonstrated superiority to education alone.[10,74–76] Researchers have attempted to determine the individual modifications that create the most change in weight. Specific attention has been paid to the difference between exercise interventions and dietary alterations. Although some differences have been detected between the diet and exercise regimes' initial effects on weight, an overview of the current findings suggests that interventions that utilize both diet and physical components demonstrate long-term efficacy.[77] For example, Savoye and colleagues[78] compared outcomes of BWL and controls in 209 child participants aged 8–16 years with BMIs >95th percentile. Participants randomized to the weight-management group attended once-weekly nutritional and behavior-modification sessions and twice-weekly exercise sessions for 6 months, after which the sessions continued biweekly for another 6 months. The control group, who was seen every 6 months, received diet and exercise counseling with registered dietitians and physicians, as well as brief counseling from a social worker. Improvements demonstrated at the 6-month mark were maintained by the end of the 12-month study. Specifically, the weight-management group showed evidence of reduced percentage of body fat, total body fat (when considering growth in height), and circulating concentrations of cholesterol; controls increased in all areas. Similar components were considered in Hughes *et al.*'s study of Scottish children.[79] When compared with standard dietetic care, individual behavioral interventions demonstrated significant benefits in physical activity and sedentary behavior, suggesting that the inclusion of these behavioral-change targets is worthwhile in the treatment of pediatric obesity.

Despite such findings, questions about the efficacy of BWL still remain. For instance, one meta-analysis of behavioral interventions examined randomized controlled trials of children and adolescents aged 2–18 years. Specifically, it considered interventions such as information/education and behavioral, cognitive, social-support, and environmental strategies aimed at changing lifestyle behaviors, such as activity levels and dietary habits. Overall, the interventions exhibited small changes in target behaviors, with no significant effect on BMI when compared with controls.[80] Examining these results further, it appears that trials of intervention periods >6 months demonstrated a larger reduction in sedentary activity and BMI.[80] Despite their insignificant effect on BMI, these findings suggest that more protracted interventions that utilize longer time frames for outcome measurements may be more effective in creating observable change in both BMI and healthy behaviors. Results from this study also found a difference between child and adolescent trials, with children demonstrating larger reductions in sedentary activity. The variable effects of intervention strategies may be a function of the specific stage of a child's development. This issue requires continued attention through research, because 25% to 50% of obese adolescents remain obese in adulthood.[81–83] Alternate interventions, some of which attempt to address the unique issues of change in adolescents, are described below.

Family-Based Treatment

As discussed above, previous research has demonstrated in some instances that behavioral treatment programs applied to obese adolescents are less successful than when applied to children.[58,73] In light of this, some argue that adolescence may be too late of an intervention point; a potential advantage of addressing obesity in childhood is that the parental influence is still strong and obesity-promoting behaviors have yet to become well-engrained.[84] Moreover, children, perhaps even more so than adolescents, are embedded in multiple environments (ie, home environment, school environment) where they are vulnerable to others' decisions regarding the accessibility and availability of healthy food, exercise, and overall health-promoting lifestyles. Parents are therefore an important avenue of intervention, as they exert significant amount of influence over the home environment.

In family-based interventions, parents are viewed as an integral part of the child's treatment. The interventions cannot occur without the consent of 1 or more parents to fully participate, as parents are expected to demonstrate a significant involvement in the process of change. Strategies such as goal setting, reinforcement, contingencies, environmental control, and self-monitoring are still used. Parents are often viewed as a model for their children, and are asked to set goals for themselves as well as to model healthy changes in eating and physical activity expected of their child or adolescent. Some family-based treatments involve overweight parents in attempts at personal weight loss as well, encouraging this additional form of modeling for their child, as well as increasing hopes of maintaining a healthy home environment. Additionally, attempts are made to enhance parenting skills through parent training.

Some approaches place a large emphasis on the importance of parent training. Research interested in the effect of the parental component[85] examined the effectiveness of parenting-skills training complemented by lifestyle education in a sample of 111 Australian children. They found that both parenting-skills training and lifestyle education are important components in addressing childhood obesity. Moreover, Golan and colleagues[86–88] assert that parents may be such an integral force in their child's weight loss that the child is not even necessary in sessions. Instead, sessions address parenting skills, with specific focus on an authoritative parenting style. In these sessions, much like other family-based treatments, parents receive nutritional education and discuss eating and activity behavior modification, decreasing stimulus exposure, parental modeling, problem solving, and strategies for coping with resistance. Seven-year follow-up data of 50 children (aged 14–19 years at follow-up) whose parents participated in health-centered family-based interventions showed a 29% reduction in overweight, greater than in child-only interventions. Here, parents were the sole participants in group sessions, though changes implemented were intended for the entire family.[88]

Epstein *et al.*'s study of overweight children has become a cornerstone in the support for family-based treatment.[73,84,89–93] In this body of research, overweight children aged 6–12 years were treated with a family-based intervention that focused on creating behavior change in the areas of diet, physical activity, and sedentary behaviors. Children in the intervention group experienced a decrease of 11.2% in overweight after 5 years and a 7.5% decrease after 10 years, compared with increases of 7.9% and 14.3% in the control condition. Further, when parents' weight was also targeted, decreases in baseline weights for children were reported at −12% (compared with +4% in the control condition) at 5 years and −15% (compared with +3%

in the control condition) at 10 years.[90] Interested in examining the enduring efficacy of this approach, Epstein *et al.*[93] compared the efficacy of interventions initiated >20 years ago to current programs through a 24-month follow-up period. Results of this comparison showed significant reductions over time in BMI z-score and percent overweight in both contemporary and older intervention studies, suggesting the continued effectiveness of this approach.

Further support for the familial component of weight management is presented by Kitzmann and Beech in their review of 31 family-based intervention studies, in which parents participated in behavioral or cognitive behavioral approaches altering eating and exercising behaviors.[94] The authors were specifically interested in looking at the level and variety of parent involvement in family-based treatment. Results from their meta-analysis suggest that the addition of parental involvement, regardless of type or level, consistently demonstrates some change in child weight when compared with control groups.

A meta-analysis of 31 studies involving family-based interventions in weight management suggests that the addition of parental involvement, regardless of type or level, consistently demonstrates some change in child weight when compared with control groups.

Despite these findings, some issues still remain. Much of the research for family-based interventions has been done with moderately obese children; however, a more severely obese population may challenge intervention efficacy. Interested in addressing this concern, Kalarchian *et al.*[95] examined the effect of family-based interventions on a sample of 192 children aged 8–12 years with a BMI ≥97th percentile, adapting the study intervention developed by Epstein *et al.*[73,75,89–92] Although they received complementary material, adult and child groups were seen separately for 20 group meetings over the course of 6 months. Six booster sessions were provided between month 6 and month 12. Follow-up assessments were taken at month 12 and month 18. Control groups were offered 2 nutritional consultation sessions to develop individual nutrition plans. At the 6-month mark, when compared with controls, family-based treatment was associated with significant levels of decrease in percent overweight. After 12 months, family-based participants sustained improvements in medical risk parameters such as waist circumference, systolic blood pressure, percent body fat, and total body fat.[95] However, the changes demonstrated in percent overweight were not well maintained by the family-based participants past the 6-month period. Instead, percent overweight and BMI of these participants steadily increased after weekly intervention sessions ended. Conversely, control group participants maintained a stable overweight status, yielding little contrast between the 2 groups at 12-month and 18-month follow-ups.[56] These findings are somewhat surprising, as previous studies have suggested long-term efficacy of family-based interventions in pediatric obesity. In this way, it appears that maintenance of weight loss progress in severely obese pediatric populations may present increased challenges. Future research should consider this challenging population, which is increasing over time, when determining efficacy of interventions.

Research also exhibits little evidence to support family-based interventions in adolescents, because overall the literature targeting adolescents is not as well developed as the literature for children. This is problematic because the simple application of family-based interventions to an older population suggests a one-size-fits-all approach, when the fundamental development of these age groups, and corresponding changes in the appropriate quality and quantity of parental involvement, are potentially significantly different. Though there is no official, empirically supported treatment yet that proves developmentally sensitive for adolescents, researchers have begun to explore these issues, which are touched on below.

Current Innovation in the Prevention and Treatment of Adolescent Obesity

Researchers at the Mount Sinai School of Medicine and the University of Chicago, interested in exploring alternative techniques for combating pediatric obesity, are examining the efficacy of an adapted eating disorder intervention for obese adolescents (1R21HD057394-01A1; PI: Loeb). Eating disorder literature has supported the use of the manualized Maudsley-based family therapy, which considers parents an integral part of weight restoration, for the treatment of anorexia nervosa and bulimia nervosa.[96–98] The broader adapted intervention for obesity[99] is designed to be transdevelopmental, with separate modules for early childhood, middle childhood, and adolescence (modulating the nature and intensity of parental involvement as a function of developmental stage); the current study is testing the adolescent module in particular. Currently treatment of this nature is purely investigatory, but it is still critical to consider its potential implications.

Treatment involves 16 sessions completed over the course of 24 weeks. Sessions are divided into 3 phases, with parents having more active involvement in the early stages of treatment and the offspring gaining more autonomy as the phases and corresponding behavioral changes progress. An emphasis is placed on parental strategies, with parents expected to serve both as models and guides. Interventions incorporate some behavioral strategies, such as self-monitoring, goal setting, environmental control, and reinforcement. An equally important aspect of this intervention is the reduction of blame and externalization of weight struggles. Because frequent failed attempts at weight loss, increased health problems, and overall dissatisfaction can lead to a high level of frustration, this intervention attempts to safeguard against increased negativity. Here, specific focus is placed on externalizing the weight problems, separating the obesity from the child. In this way the weight is framed as an outside factor afflicting the child, thus reducing blame of both the child and the parents, as obesity is not believed to be "caused" by either. This reduction in blame and externalization of the problem then allows the parents to attack the weight without attacking the child.

Recently, Stice and colleagues[100–104] have provided evidence of successful interventions for the onset of obesity in 481 adolescent girls (aged 14–19 years) who internalized the "thin ideal." Participants were randomized to a dissonance intervention, a healthy weight intervention, an expressive-writing control intervention, or an assessment-only control condition. The dissonance intervention required adolescents to engage in verbal, written, and behavioral exercises critiquing the thin ideal in order to create psychological discomfort and reduce internalization. This intervention strategy draws on the tenets of cognitive dissonance theory, in which an individual consistently acts as though he follows a specified set of beliefs, with the understanding that eventually the belief, or mindset, catches up to the behavior, creating a change in both belief and behavior. The healthy weight intervention asked adolescents to make gradual healthy changes to their diet and physical activity using some behavior-modification techniques. The expressive-writing control intervention required adolescents to write for 45 minutes 3 times a week, and the assessment-only control did not receive any intervention during the study period.

Through a series of trials, the researchers found that adolescents receiving the dissonance intervention demonstrated significantly greater decreases in thin ideal internalization, body dissatisfaction, and psychosocial impairment than did the expressive-writing controls. Similar results were demonstrated when comparing dissonance intervention with the assessment-only controls. Adolescents in the healthy weight intervention also exhibited similar effects, with greater decreases in thin ideal internalization and weight than the expressive-writing controls, as well as lower risk for eating pathology and obesity in the following years than for assessment-only controls. Furthermore, adolescents in the dissonance intervention and healthy weight intervention produced a 55% reduction in risk for the onset of obesity relative to control groups in a 3-year follow-up period. Such findings suggest the more enduring effect of interventions than previously considered, as meta-analyses have remained skeptical of long-term effects of obesity prevention programs.[103] Moreover, the implications of these findings are that both the modification of healthy eating, as well as the modification of beliefs that may trigger unhealthy dieting, affect long-term weight trajectory for adolescents. Theoretically, in this way, those who are able to reject the thin ideal may also be more able to adaptively use the healthy aspects of dieting while rejecting unhealthy aspects. It appears that obesity may be thwarted without specifically targeting weight or diet, but instead through addressing the thin ideal. The healthy weight intervention suggests that there may be a way to channel the more positive aspects of dieting without increasing the unhealthy consequences associated with extreme dieting behavior.

BINGE EATING DISORDER

Patients with BED are a specific subgroup within the obese population whose identification and treatment have been the source of much debate. BED is generally defined by the presence of binges, or eating excessive amounts of food while experiencing a loss of control, unaccompanied by regular purging with laxatives, vomiting, or excessive exercise (American Psychiatric Association, 2000). The exact relationship between BED and obesity is unclear[105] and beyond the scope of this paper. However, it is important to note that most reviews of psychological treatments for obesity include studies on both obese binge eaters and obese non–binge eaters; however, the decision was to specifically distinguish research for the 2 groups in this review. Studies have indicated high levels of disordered eating behaviors and cognitions among obese community samples,[106,107] and it seems that BED has a higher prevalence in obese populations (2.9%) compared with general populations (1.5%).[28,108] The issue then is of the appropriate treatments to use with this specialty population, who may not always be immediately identified upon entering weight-loss treatment.

> *Studies have indicated high levels of disordered eating behaviors and cognitions among obese community samples, and it seems that binge eating disorder has a higher prevalence in obese populations (2.9%) compared with general populations (1.5%).*

Early studies comparing different types of BWL on obese binge-eating populations found some positive results for short-term weight loss,[109–115] although these losses were often not maintained over longer periods of time. However, in studies comparing BWL for binge eaters and non–binge eaters, binge eaters tended to lose less weight post-treatment and have higher dropout rates.[113–116] In a sample of 131 obese patients enrolled in BWL, patients with BED were significantly less likely to achieve the weight loss goals of non-BED patients (16% versus 37%).[116] A meta-analysis on the impact of BED on weight loss treatment effects (including pharmaceutical and surgical interventions) found that, on average, BED samples lost significantly less weight than the non-BED obese (−1.3 kg versus −10.5 kg).[117] It may be that response to BWL might be associated with degree of impairment; there has been some suggestion that BED patients with high shape and weight concerns have different psychopathology profiles and levels of impairment than BED cases without issues of shape and restraint.[118] When patients have been differentiated by severity of symptoms, it does seem that whereas moderate BED responds relatively well to BWL, it is less effective for more symptomatic patients.[110,114] For example, in a sample of 44 overweight and obese women, enrollment in either standard BWL or BWL modified to address binge eating did not lead to significant differences in weight loss for those with no or moderate binge eating.[110] However, those with severe binge eating saw significant weight loss when enrolled in the specialized BWL, as opposed to no change when enrolled in BWL. It seems that patients on the more severe end of the binge eating spectrum may benefit from more eating disorder–appropriate treatment,[110,119–121] although the primary benefit in these cases is reduction of binge frequency and specific eating disorder or general psychopathology, rather than weight loss.

These specialized treatments include CBT; interpersonal therapy (IPT), which addresses the connection between interpersonal relationships and the development of psychological symptoms; and dialectical behavior therapy (DBT), a therapy originally developed for patients with borderline-personality disorder that combines CBT techniques with Eastern principles of mindfulness and self-acceptance. Results on CBT and IPT have been the most promising. CBT is generally found to be more effective than BWL[40,120,121] and some research has found guided self-help CBT (CBT-GSH) to be effective in reducing clinical severity compared with wait-list controls or BWL, which is promising given its cost-effectiveness and easy dissemination.[52,54,122] Studies comparing CBT and IPT have been inconclusive,[10,123] although a recent study comparing CBT-GSH and IPT concluded that CBT-GSH was effective for all but the most severe BED patients, wherein IPT was the recommended alternative.[55] Available data on DBT is limited, although the existing literature shows some promise.[11,124,125] It seems that in general, the issue of whether BED is an obesity subtype or simply a marker for comorbid psychopathology needs to be resolved[105,126] in order to determine which treatments are the most effective. It is important to note that the efficacy of existing treatments for BED rests upon their effect in reducing eating-disorder and general psychology and binge frequency, rather than facilitating weight loss.[105,108] It remains to be determined if weight loss can reasonably be achieved within this specific obese population.

Recent research suggests that some level of binge-eating behavior may also occur in the pediatric overweight and obese population as well.[127,128] This area is still new to the adolescent weight-management field, yet there appears to be some important implications for its development. Morgan and colleagues[129] found that children aged 6–10 years who reported experiencing episodes of loss of control over eating weighed more and had greater amounts of body fat than children who did not report a loss of control. As supported in previous research,[130] loss-of-control children also endorsed higher levels of anxiety, depressive symptoms, and body dissatisfaction. In a study of 160 overweight adolescents (BMI 40.7 ± 8.8 kg/m^2), those reporting symptoms of full BED also reported higher scores on the Eating Disorder Examination, as well as higher levels of anxiety and more negative mood than adolescents exhibiting fewer binge behaviors.[131] Furthermore, adolescents who endorsed having recent (within the last 3 months) but infrequent binge behavior reported higher scores on the Eating Disorder Examination and higher levels of anxiety than adolescents without a history of binge behaviors. This suggests the significant consequence of subthreshold BED behaviors and implicates the need for continual attention and incorporation of binge-eating behavior in the conceptualization of pediatric overweight and obesity interventions.

ADVANCES IN DISSEMINATING INTERVENTIONS

The challenge of providing effective and manageable treatment for overweight and obesity has prompted researchers to explore of a variety of avenues of treatment. One such avenue has been the use of the Internet to implement behavior change. Although computer use may seem almost intuitive as the computer has become ubiquitous in homes, it is important to recognize that some limitations remain. Health-information consumers on the Internet still tend to be white, non-Hispanic, younger, and better educated, with higher income and from urban areas.[132]

Web-based interventions may be used in conjunction with the aforementioned approaches to weight loss. Typically, they aim to promote behavior change toward a healthy lifestyle by providing online health-education materials related to food, physical activity, weight loss, and body image, as well as implementing strategies such as self-monitoring, goal setting, problem solving, and behavioral contracting.[133] Web-based interactive programs including games, quizzes, and charts to plan daily activities aim to sustain attention while teaching new strategies for healthy living. Some interventions require participants to sign on to a secure Web site where behavioral information is logged; others send CD-ROMs to participants over set time intervals.[134–136] Treatment approaches have been applied to the child specifically or in family-based interventions involving the participation of both parents and child.[137–139]

Web-based interventions were examined in an ethnically diverse population of overweight children aged 12–18 years with a BMI of 34.2 ± 7.0 kg/m^2. Participants receiving an Internet-delivered cognitive behavioral program including psycho-educational readings, online asynchronous discussion groups, online self-monitoring of food/activity/weight and body image, weekly e-mail newsletters containing individualized feedback, and telephone calls demonstrated a significant reduction in BMI. Further, when compared with participants who received only informational handouts, the Web-based participants demonstrated significantly increased healthy eating–related and physical activity–related cognitive behavioral skill.[134] Further review of Web-based interventions suggests similar results, with 3 studies of overweight African American adolescents (aged 15–17 years) demonstrating comparably more weight loss in Web-based groups compared with controls.[133] As mentioned, some Web-based interventions involved both the parent and the child; An *et al.*[133] found that family-based parental involvement is effective in child weight loss and may even increase likelihood of success in weight loss compared with Web-based interventions without parental support. This echoes researchers' arguments for the utility of family-based treatment, as familial involvement, even Web-based, demonstrates successful weight loss outcomes.

The literature on Web-based weight-loss treatments for adults is less positive, although there is some evidence for increased efficacy given further development and refinement of current online interventions. Harvey-Berino *et al.*[140,141] found no significant differences in long-term weight maintenance between groups randomized to Internet support versus frequent and infrequent in-person contact, whereas Wing *et al.*[29] reported a significant difference in the proportion of post-treatment patients who gained weight in an information-only group ($n = 105$; 72.4%) compared with Internet-based support ($n = 104$; 54.8%) or face-to-face contact ($n = 105$; 45.7%), although only significant differences in actual weight regain were observed between the face-to-face and control groups (2.5 ± 6.7 kg versus 4.9 ± 6.5 kg). However, other studies have found Internet-based interventions to be much less effective than face-to-face contact or a behavioral weight-loss manual.[142,143] Some of these effects may be mediated by frequency of use, where responders seem to show significantly more weight loss than noncompliers.[143–146] It therefore seems important to identify the specific components of a successful weight loss site. Krukowski *et al.*[147] surveyed participants in a 12-month behavioral weight loss program and found that during the initial treatment, utilization of feedback factors such as progress reports and physiological calculators predicted weight loss. Social-support features, such as message boards or online chats, did not demonstrate a significant effect during the weight loss period,[144,145] although they may have more efficacy for long-term weight maintenance.[147]

Internet interventions for adults certainly hold promise; reviews of Web-based treatments for health promotion and weight loss have been mixed, but also clearly identify some evidence for effectiveness compared with controls.[148–150] The accessibility of the Internet to the general public makes it the best tool to engage large sectors of the target population, as shown by Petersen *et al.*'s Internet-based employee weight-management program, which enrolled more than 7000 participants, although actual effects were minimal.[151] For the long term, it is also much more cost-effective than traditional face-to-face treatments, once fixed set-up costs have been mitigated.[144] Future research needs to incorporate elements of successful existing programs and also evaluate outcomes at longer follow-up periods to determine long-term efficacy and address crucial factors such as high attrition rates and noncompliance.

CONCLUSION

It seems that psychological treatments have an important role to play in weight loss and weight-maintenance interventions. Besides improving general quality of life and achieving modest weight loss, they can play an important role in maintaining weight losses.

There are several methodological limitations that make it difficult to accurately judge the effectiveness of different psychological treatments across clinical trials. There is no comprehensive validated treatment for weight loss interventions,[15] which leads to variability across studies in terms of exact format, specific dietary and exercise goals, frequency of sessions, delivery of strategies, and provider licensure. Particularly for psychological treatments, there does not seem to be consistent literature on the various treatments, as most studies use different variations or combinations of treatment types,[14] which makes generalizations difficult. While the National Heart, Lung, and Blood Institute recommended in 2005 that weight loss studies have at least a 5-year follow-up period, few studies report data for more than 3 years. Generalizability issues are also a concern, as most research to date has been conducted in university settings. Commercial weight loss programs, which were not explored in this review, are a possible area of dissemination. Dissemination is a major concern that needs to be addressed; Bish *et al.*[152] reported that although in 2000 46% of women and 33% of men reported trying to lose weight, less than 10% actually followed minimum recommendations for weight loss. As a whole, these limitations make it difficult to draw concrete conclusions on the long-term efficacy of psychological treatments, underlining the need for further research focused on this issue.

In synthesizing the available literature on current research it is clear that the field of psychological and behavioral interventions has made significant strides; however, there is still much room for growth. For pediatric obesity, interventions are required that specifically demonstrate increased sensitivity to the developmental needs of adolescence. Moreover, it is essential that research turn its attention to the more severely obese youth, as this increasing population is in significant need of intervention strategies that will increase the likelihood of long-term weight-loss maintenance. Weight loss treatments for adults should continue to focus on the issue of long-term weight maintenance, placing emphasis on exploring the role of cognitive behavioral treatment and the role of the Internet. Further research is also necessary in distinguishing obese subpopulations, such as those with BED, in order to determine specialty interventions that might be more efficacious in reducing associated distress along with promoting weight loss.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Haslam DW, James WP. Obesity. *Lancet* 2005; 366: 1197–1209.
2. Popkin BM, Doak CM. The obesity epidemic is a worldwide phenomenon. *Nutr Rev* 1998; 56: 106–114.
3. Wang Y, Beydoun MA. The obesity epidemic in the United States–gender, age, socioeconomic, racial/ethnic, and geographic characteristics: a systematic review and meta-regression analysis. *Epidemiol Rev* 2007; 29: 6–28.
4. Flegal KM, Carroll MD, Kuczmarski RJ, Johnson CL. Overweight and obesity in the United States: prevalence and trends, 1960–1994. *Int J Obes Relat Metab Disord* 1998; 22: 39–47.
5. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999–2008. *JAMA* 2010; 303: 235–241.
6. Overweight, obesity, and health risk: National Task Force on Prevention and Treatment of Diabetes. *Arch Intern Med* 2000; 160: 898–904.
7. Field AE, Coakley EH, Must A, et al. Impact of overweight on the risk of developing common chronic diseases during a 10-year period. *Arch Intern Med* 2001; 161: 1581–1586.
8. Kushner RF, Foster GD. Obesity and quality of life. *Nutrition* 2000; 16: 947–952.

9. De Zwaan M, Petersen I, Kaerber M, et al. Obesity and quality of life: a controlled study of normal-weight and obese individuals. *Psychosomatics* 2009; 50: 474–482.
10. Wilfley DE, Welch RR, Stein RI, et al. A randomized comparison of group cognitive-behavioral therapy and group interpersonal psychotherapy for the treatment of overweight individuals with binge-eating disorder. *Arch Gen Psychiatry* 2002; 59: 713–721.
11. Telch CF, Agras WS, Linehan MM. Dialectical behavior therapy for binge eating disorder. *J Consult Clin Psychol* 2001; 69: 1061–1065.
12. Douketis JD, Macie C, Thabane L, Williamson DF. Systematic review of long-term weight loss studies in obese adults: clinical significance and applicability to clinical practice. *Int J Obes (Lond)* 2005; 29: 1153–1167.
13. Hankey CR. Session 3 (Joint with the British Dietetic Association): management of obesity: weight-loss interventions in the treatment of obesity. *Proc Nutr Soc* 2010; 69: 34–38.
14. Shaw K, O'Rourke P, Del Mar C, Kenardy J. Psychological interventions for overweight or obesity. *Cochrane Database Syst Rev* 2005; 18: CD003818.
15. Paul-Ebhohimhen V, Avenell A. A systematic review of the effectiveness of group versus individual treatments for adult obesity. *Obes Facts* 2009; 2: 17–24.
16. Jeffery RW, Drewnowski A, Epstein LH, et al. Long-term maintenance of weight loss: current status. *Health Psychol* 2000; 19(suppl 1): 5–16.
17. Clifton PM. Dietary treatment for obesity. *Nat Clin Pract Gastroenterol Hepatol* 2008; 5: 672–681.
18. Israel AC, Saccone AJ. Follow-up of effects of choice of mediator and target of reinforcement on weight loss. *Behav Ther* 1979; 10: 260–265.
19. Gormally J. Rardin D. Weight loss and maintenance and changes in diet and exercise for behavioral counseling and nutrition education. *J Couns Psychol* 1981; 28: 295–304.
20. Wing RR, Epstein LH, Marcus MD, Koeske R. Intermittent low-calorie regimen and booster sessions in the treatment of obesity. *Behav Res Ther* 1984; 22: 445–449.
21. Jeffery RW, Wing RR. Long-term effects of interventions for weight loss using food provision and monetary incentives. *J Consult Clin Psychol* 1995; 63: 793–796.
22. Wadden TA, Butryn ML, Wilson C. Lifestyle modification for the management of obesity [published correction appears in *Gastroenterology* 2007;133:371]. *Gastroenterology* 2007; 132: 2226–2238.
23. Perri MG, Fuller PR. Success and failure in the treatment of obesity: where do we go from here? *Med Exer Nutr Health* 1995; 4: 255–272.
24. Kramer FM, Jeffery RW, Forster JL, Snell MK. Long-term follow-up of behavioral treatment for obesity: patterns of weight regain among men and women. *Int J Obes* 1989; 13: 123–136.
25. Stevens VJ, Obarzanek E, Cook NR, et al. Trials of Hypertension Prevention Research Group. Long-term weight loss and changes in blood pressure: results of the Trials of Hypertension Prevention, phase II. *Ann Intern Med* 2001; 134: 1–11.
26. Jeffery RW, Drewnowski A, Epstein LH, et al. Long-term maintenance of weight loss: current status.. *Health Psychol* 2000; 19(suppl 1): 5–16.
27. Perri MG, Nezu AM, McKelvey WF, et al. Relapse prevention training and problem-solving therapy in the long-term management of obesity. *J Consult Clin Psychol* 2001; 69: 722–726.
28. Carels RA, Darby L, Cacciapaglia HM, et al. Applying a stepped-care approach to the treatment of obesity. *J Psychosom Res* 2005; 59: 375–383.
29. Wing RR, Tate DF, Gorin AA, et al. A self-regulation program for maintenance of weight loss. *N Engl J Med* 2006; 355: 1563–1571.
30. Svetkey LP, Stevens VJ, Brantley PJ, et al.; Weight Loss Maintenance Collaborative Research Group. Comparison of strategies for sustaining weight loss: the Weight Loss Maintenance Randomized Controlled Trial. *Am J Health Promot* 2008; 23: 152–153.
31. Perri MG, Limacher MC, Durning PE, et al. Extended-care programs for weight management in rural communities: the Treatment of Obesity in Underserved Rural Settings (TOURS) randomized trial. *Arch Intern Med* 2008; 168: 2347–2354.
32. Jakicic JM, Marcus BH, Lang W, Janney C. Effect of exercise on 24-month weight loss maintenance in overweight women [published correction to *Arch Intern Med* 2008;168:1550–1560]. *Arch Intern Med* 2008; 168: 2162.
33. Wing RR, Jeffery RW. Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. *J Consult Clin Psychol* 1999; 67: 132–138.
34. Gorin A, Phelan S, Tate D, et al. Involving support partners in obesity treatment. *J Consult Clin Psychol* 2005; 73: 341–343.
35. Jeffery RW, Wing RR, Thorson C, et al. Strengthening behavioral interventions for weight loss: a randomized trial of food provision and monetary incentives. *J Consult Clin Psychol* 1993; 61: 1038–1045.
36. The Diabetes Prevention Program. Design and methods for a clinical trial in the prevention of type 2 diabetes [published correction appears in *Diabetes Care* 1999;22:1389]. *Diabetes Care* 1999; 22: 623–634.
37. Ryan DH, Espeland MA, Foster GD, et al. Look AHEAD (Action for Health in Diabetes): design and methods for a clinical trial of weight loss for the prevention of cardiovascular disease in type 2 diabetes. *Control Clin Trials* 2003; 24: 610–628.
38. Espeland MA, Bray GA, Neiberg R, et al. Describing patterns of weight changes using principal components analysis: results from the Action for Health in Diabetes (Look AHEAD) Research Group. *Ann Epidemiol* 2009; 19: 701–710.
39. Hamman RF, Wing RR, Edelstein SL, et al. Effect of weight loss with lifestyle intervention on risk of diabetes. *Diabetes Care* 2006; 29: 2102–2107.
40. Knowler WC, Fowler SC, Hamman RF, et al. Diabetes Prevention Program Research Group. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study [published correction appears in *Lancet* 2009;374:2054]. *Lancet* 2009; 374: 1677–1686.
41. Liao KL. Cognitive-behavioural approaches and weight management: an overview. *J R Soc Promot Health* 2000; 120: 27–30.
42. Cooper Z, Fairburn CG. A new cognitive behavioural approach to the treatment of obesity. *Behav Res Ther* 2001; 39: 499–511.

43. Van Dorsten B, Lindley EM. Cognitive and behavioral approaches in the treatment of obesity. *Endocrinol Metab Clin North Am* 2008; 37: 905–922.
44. Block JD. Effects of rational-emotive therapy on overweight adults. *Psychother: Theory Res Pract* 1980; 17: 277–280.
45. Dennis KE, Pane KW, Adams BK, Qi BB. The impact of a shipboard weight control program. *Obes Res* 1999; 7: 60–67.
46. Rodriguez-Hernandez H, Morales-Amaya UA, Rosales-Valdez R, et al. Adding cognitive behavioural treatment to either low-carbohydrate or low-fat diets: differential short-term effects. *Br J Nutr* 2009; 102: 1847–1853.
47. Werrij MQ, Jansen A, Mulkens S, et al. Adding cognitive therapy to dietetic treatment is associated with less relapse in obesity. *J Psychosom Res* 2009; 67: 315–324.
48. Painot D, Jotterand S, Kammer A, et al. Simultaneous nutritional cognitive-behavioural therapy in obese patients. *Patient Educ Couns* 2001; 42: 47–52.
49. Sbrocco T, Nedegaard RC, Stone JM, Lewis EL. Behavioral choice treatment promotes continuing weight loss: preliminary results of a cognitive-behavioral decision-based treatment for obesity. *J Consult Clin Psychol* 1999; 67: 260–266.
50. Fabricatore AN. Behavior therapy and cognitive-behavioral therapy of obesity: is there a difference? *J Am Diet Assoc* 2007; 107: 92–99.
51. Rapoport L, Clark M, Wardle J. Evaluation of a modified cognitive-behavioural programme for weight management. *Int J Obes Relat Metab Disord* 2000; 24: 1726–1737.
52. Carter JC, Fairburn CG. Cognitive-behavioral self-help for binge eating disorder: a controlled effectiveness study. *J Consult Clin Psychol* 1998; 66: 616–623.
53. Loeb KL, Wilson GT, Gilbert JS, Labouvie E. Guided and unguided self-help for binge eating. *Behav Res Ther* 2000; 38: 259–272.
54. Peterson CB, Mitchell JE, Engbloom S, et al. Self-help versus therapist-led group cognitive-behavioral treatment of binge eating disorder at follow-up. *Int J Eat Disord* 2001; 30: 363–374.
55. Wilson GT, Wilfley DE, Agras WS, Bryson SW. Psychological treatments of binge eating disorder. *Arch Gen Psychiatry* 2010; 67: 94–101.
56. McGuire MT, Wing RR, Klem ML, et al. Long-term maintenance of weight loss: do people who lose weight through various weight loss methods use different behaviors to maintain their weight? *Int J Obes Relat Metab Disord* 1998; 22: 572–577.
57. Wing RR, Hill JO. Successful weight loss maintenance. *Annu Rev Nutr* 2001; 21: 323–341.
58. Brownell KD, Kelman JH, Stunkard AJ. Treatment of obese children with and without their mothers: changes in weight and blood pressure. *Pediatrics* 1983; 71: 515–523.
59. Hedley AA, Ogden CL, Johnson CL, et al. Prevalence of overweight and obesity among US children, adolescents, and adults, 1999–2002. *JAMA* 2004; 291: 2847–2450.
60. Ogden CL, Carroll MD, Flegal KM. High body mass index for age among US children and adolescents, 2003–2006. *JAMA* 2008; 299: 2401–2405.
61. Summerbell CD, Ashton V, Campbell KJ, et al. Interventions for treating obesity in children. *Cochrane Database Syst Rev* 2003; 3: CD001872.
62. Dietz WH. Health consequences of obesity in youth: childhood predictors of adult disease. *Pediatrics* 1998; 101: 518–525.
63. Eisenberg ME, Neumark-Sztainer D, Story M. Associations of weight-based teasing and emotional well-being among adolescents. *Arch Pediatr Adolesc Med* 2003; 157: 733–738.
64. Freedman DS, Dietz WH, Srinivasan SR, Berenson GS. The relation of overweight to cardiovascular risk factors among children and adolescents: the Bogalusa Heart Study. *Pediatrics* 1999; 103: 1175–1182.
65. Latner JD, Stunkard AJ, Wilson GT, Jackson ML. The perceived effectiveness of continuing care and group support in the long-term self-help treatment of obesity. *Obesity (Silver Spring)* 2006; 14: 464–471.
66. Srinivasan SR, Myers L, Berenson GS. Predictability of childhood adiposity and insulin for developing insulin resistance syndrome (syndrome X) in young adulthood: the Bogalusa Heart Study. *Diabetes* 2002; 51: 204–209.
67. Strauss RS, Pollack HA. Social marginalization of overweight children. *Arch Pediatr Adolesc Med* 2003; 157: 746–752.
68. Dabelea D, Knowler WC, Pettitt DJ. Effect of diabetes in pregnancy on offspring: follow-up research in the Pima Indians. *J Matern Fetal Med* 2000; 9: 83–88.
69. Glenny AM, O'Meara S, Melville A, et al. The treatment and prevention of obesity: a systematic review of the literature. *Int J Obes Relat Metab Disord* 1997; 21: 715–737.
70. McGovern L, Johnson JN, Paulo R, et al. Treatment of pediatric obesity: a systematic review and meta-analysis of randomized trials. *J Clin Endocrinol Metab* 2008; 93: 4600–4605.
71. Zametkin AJ, Zoon CK, Klein HW, Munson S. Psychiatric aspects of child and adolescent obesity: a review of the past 10 years. *J Am Acad Child Adolesc Psychiatry* 2004; 43: 134–150.
72. Oude Luttikhuis H, Baur L, Jansen H, et al. Interventions for treating obesity in children [update of *Cochrane Database Syst Rev* 2003;3:CD001872]. *Cochrane Database Syst Rev* 2009; 1: CD001872.
73. Epstein LH, Wing RR. Behavioral treatment of childhood obesity [published correction appears in *Psychol Bull* 1988;103:323]. *Psychol Bull* 1987; 101: 331–342.
74. Saelens BE, Sallis JF, Wilfley DE, et al. Behavioral weight control for overweight adolescents initiated in primary care. *Obes Res* 2002; 10: 22–32.
75. Epstein LH, Wing RR, Steranchak L, et al. Comparison of family based behavior modification and nutrition education for childhood obesity. *J Pediatr Psychol* 1980; 5: 25–36.
76. Israel AC, Stolmaker L, Sharp JP, et al. An evaluation of two methods of parental involvement in treating obese children. *Behav Ther* 1984; 15: 266–272.
77. Jelalian E, Saelens BE. Empirically supported treatments in pediatric psychology: pediatric obesity. *J Pediatr Psychol* 1999; 24: 223–248.
78. Savoye M, Shaw M, Dziura J, et al. Effects of a weight management program on body composition and metabolic parameters in overweight children:

a randomized controlled trial. *JAMA* 2007; 297: 2697–2704.

79. Hughes AR, Stewart L, Chapple J, et al. Randomized, controlled trial of a best-practice individualized behavioral program for treatment of childhood overweight: Scottish Childhood Overweight Treatment Trial (SCOTT). *Pediatrics* 2008; 121: e539–e546.
80. Kamath CC, Vickers KS, Ehrlich A, et al. Behavioral interventions to prevent childhood obesity: a systematic review and meta-analyses of randomized trials. *J Clin Endocrinol Metab* 2008; 93: 4606–4615.
81. Charney E, Goodman HC, McBride M, et al. Childhood antecedents of adult obesity: do chubby infants become obese adults? *N Engl J Med* 1976; 295: 6–9.
82. Guo SS, Roche AF, Chumlea WC, et al. The predictive value of childhood body-mass index values for overweight at age 35 y. *Am J Clin Nutr* 1994; 59: 810–819.
83. Garn S. Continuities and changes in fatness from infancy through adulthood. *Curr Probl Pediatr* 1985; 15: 1–47.
84. Epstein LH, Myers MD, Raynor HA, Saelens BE. Treatment of pediatric obesity. *Pediatrics* 1998; 101: 554–570.
85. Golley RK, Magarey AM, Baur LA, et al. Twelve-month effectiveness of a parent-led, family-focused weight-management program for prepubertal children: a randomized, controlled trial. *Pediatrics* 2007; 119: 517–525.
86. Golan M, Weizman A. Familial approach to the treatment of childhood obesity: conceptual model. *J Nutr Educ* 2001; 33: 102–107.
87. Golan M, Weizman A, Apter A, Fainaru M. Parents as the exclusive agents of change in the treatment of childhood obesity. *Am J Clin Nutr* 1998; 67: 1130–1135.
88. Golan M, Crow S. Targeting parents exclusively in the treatment of childhood obesity: long-term results. *Obes Res* 2004; 12: 357–361.
89. Epstein LH, McCurley J, Wing RR, Valoski A. Five-year follow-up of family-based behavioral treatments for childhood obesity. *J Consult Clin Psychol* 1990; 58: 661–664.
90. Epstein LH, Valoski A, Wing RR, McCurley J. Ten-year outcomes of behavioral family-based treatment for childhood obesity. *Health Psychol* 1994; 13: 373–383.
91. Epstein LH, Wing RR, Woodall K, et al. Effects of family-based behavioral treatment on obese 5- to 8-year-old children. *Behav Ther* 1985; 16: 205–212.
92. Epstein LH. Family based behavioural intervention for obese children. *Int J Obes Relat Metab Disord* 1996; 20(suppl 1): S14–S21.
93. Epstein LH, Paluch RA, Roemmich JN, Beecher MD. Family-based obesity treatment, then and now: twenty-five years of pediatric obesity treatment. *Health Psychol* 2007; 26: 381–391.
94. Kitzmann KM, Beech BM. Family-based interventions for pediatric obesity: methodological and conceptual challenges from family psychology. *J Fam Psychol* 2006; 20: 175–189.
95. Kalarchian MA, Levine MD, Arslanian SA, et al. Family-based treatment of severe pediatric obesity: randomized, controlled trial. *Pediatrics* 2009; 124: 1060–1068.
96. Eisler I, Dare C, Hodes M, et al. Family therapy for adolescent anorexia nervosa: the results of a controlled comparison of two family interventions. *J Child Psychol Psychiatry* 2000; 41: 727–736.
97. Eisler I, Dare C, Russell GF, et al. Family and individual therapy in anorexia nervosa: a 5-year follow-up. *Arch Gen Psychiatry* 1997; 54: 1025–1030.
98. Eisler I, Simic M, Russell GF, Dare C. A randomised controlled treatment trial of two forms of family therapy in adolescent anorexia nervosa: a five-year follow-up. *J Child Psychol Psychiatry* 2007; 48: 552–560.
99. Loeb KL, Celio Doyle A, Le Grange D, et al. *Family-Based Treatment for Child and Adolescent Overweight: A Transdevelopmental Approach* [unpublished manual]. New York, NY: Mount Sinai School of Medicine; 2010.
100. Stice E, Mazotti L, Weibel D, Agras WS. Dissonance prevention program decreases thin-ideal internalization, body dissatisfaction, dieting, negative affect, and bulimic symptoms: a preliminary experiment. *Int J Eat Disord* 2000; 27: 206–217.
101. Stice E, Chase A, Stormer S, Appel A. A randomized trial of a dissonance-based eating disorder prevention program. *Int J Eat Disord* 2001; 29: 247–262.
102. Stice E, Trost A, Chase A. Healthy weight control and dissonance-based eating disorder prevention programs: results from a controlled trial. *Int J Eat Disord* 2003; 33: 10–21.
103. Stice E, Shaw H, Burton E, Wade E. Dissonance and healthy weight eating disorder prevention programs: a randomized efficacy trial. *J Consult Clin Psychol* 2006; 74: 263–275.
104. Stice E, Marti CN, Spoor S, et al. Dissonance and healthy weight eating disorder prevention programs: long-term effects from a randomized efficacy trial. *J Consult Clin Psychol* 2008; 76: 329–340.
105. Wonderlich SA, Gordon KH, Mitchell JE, et al. The validity and clinical utility of binge eating disorder. *Int J Eat Disord* 2009; 42: 687–705.
106. Darby A, Hay P, Mond J, et al. Disordered eating behaviours and cognitions in young women with obesity: relationship with psychological status. *Int J Obes (Lond)* 2007; 31: 876–882.
107. Darby A, Hay P, Mond J, et al. The rising prevalence of comorbid obesity and eating disorder behaviors from 1995 to 2005. *Int J Eat Disord* 2009; 42: 104–108.
108. Striegel-Moore RH, Franko DL. Should binge eating disorder be included in the DSM-V? A critical review of the state of the evidence. *Annu Rev Clin Psychol* 2008; 4: 305–324.
109. Ho KS, Nichaman MZ, Taylor WC, et al. Binge eating disorder, retention, and dropout in an adult obesity program. *Int J Eat Disord* 1995; 18: 291–294.
110. Porzelius LK, Houston C, Smith M, et al. Comparison of a standard behavioral weight loss treatment and a binge eating weight loss treatment. *Behav Ther* 1995; 26: 119–134.
111. Gladis MM, Wadden TA, Vogt R, et al. Behavioral treatment of obese binge eaters: do they need different care? *J Psychosom Res* 1998; 44: 375–384.
112. Goodrick GK, Poston WS II, Kimball KT, et al. Nondieting versus dieting treatment for overweight binge-eating women. *J Consult Clin Psychol* 1998; 66: 363–368.

113. Raymond NC, de Zwaan M, Mitchell JE, et al. Effect of a very low calorie diet on the diagnostic category of individuals with binge eating disorder. *Int J Eat Disord* 2002; 31: 49–56.

114. Gorin A, Niemeier H, Hogan P, et al. Binge eating and weight loss outcomes in overweight individuals with type 2 diabetes: results from the Look AHEAD clinical trial. *Arch Gen Psychiatry* 2008; 65: 1447–1455.

115. Yanovski SZ, Nelson JE, Dubbert BK, Spitzer RL. Association of binge eating disorder and psychiatric comorbidity in obese subjects [published correction appears in *Am J Psychiatry* 1993;150:1910]. *Am J Psychiatry* 1993; 150: 1472–1479.

116. Pagoto S, Bodenlos JS, Kantor L, et al. Association of major depression and binge eating disorder with weight loss in a clinical setting. *Obesity* 2007; 15: 2557–2559.

117. Blaine BE, Rodman J, Newman JM. Weight loss treatment and psychological well-being: a review and meta-analysis. *J Health Psychol* 2007; 12: 66–82.

118. Grilo CM, Hrabosky JI, White MA, et al. Overvaluation of shape and weight in binge eating disorder and overweight controls: refinement of a diagnostic construct. *J Abnorm Psychol* 2008; 117: 414–419.

119. Nauta H, Hospers H, Kok G, Jansen A. A comparison between a cognitive and a behavioral treatment for obese binge eaters and obese non-binge eaters. *Behav Ther* 2000; 31: 441–461.

120. Munsch S, Biedert E, Meyer A, et al. A randomized comparison of cognitive behavioral therapy and behavioral weight loss treatment for overweight individuals with binge eating disorder. *Int J Eat Disord* 2007; 40: 102–113.

121. Wilson GT. Psychological treatment of eating disorders. *Annu Rev Clin Psychol* 2005; 1: 439–465.

122. Grilo CM, Masheb RM. A randomized controlled comparison of guided self-help cognitive behavioral therapy and behavioral weight loss for binge eating disorder. *Behav Res Ther* 2005; 43: 1509–1525.

123. Agras WS, Telch CF, Arnow B, et al. Does interpersonal therapy help patients with binge-eating disorder who fail to respond to cognitive-behavioral therapy? *J Consult Clin Psychol* 1995; 63: 356–360.

124. Wiser S, Telch CF. Dialectical behavior therapy for binge-eating disorder. *J Clin Psychol* 1999; 55: 755–768.

125. Safer DL, Robinson AH, Jo B. Outcome from a randomized controlled trial of group therapy for binge eating disorder: comparing dialectical behavior therapy adapted for binge eating to an active comparison group therapy [published correction appears in *Behav Ther* 2010;41:432]. *Behav Ther* 2010; 41: 106–120.

126. Stunkard AJ, Allison KC. Binge eating disorder: disorder or marker? *Int J Eat Disord* 2003; 34(suppl): S107–S116.

127. Decaluwe V, Braet C, Fairburn CG. Binge eating in obese children and adolescents. *Int J Eat Disord* 2003; 33: 78–84.

128. Johnson WG, Rohan KJ, Kirk AA. Prevalence and correlates of binge eating in white and African American adolescents. *Eat Behav* 2002; 3: 179–189.

129. Morgan CM, Yanovski SZ, Nguyen TT, et al. Loss of control over eating, adiposity, and psychopathology in overweight children. *Int J Eat Disord* 2002; 31: 430–441.

130. Tanofsky-Kraff M, Faden D, Yanovski SZ, et al. The perceived onset of dieting and loss of control eating behaviors in overweight children. *Int J Eat Disord* 2005; 38: 112–122.

131. Glasofer DR, Tanofsky-Kraff M, Eddy KT, et al. Binge eating in overweight treatment-seeking adolescents. *J Pediatr Psychol* 2007; 32: 95–105.

132. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. *Soc Sci Med* 2004; 59: 1795–1806.

133. An JY, Hayman LL, Park YS, et al. Web-based weight management programs for children and adolescents: a systematic review of randomized controlled trial studies. *ANS Adv Nurs Sci* 2009; 32: 222–240.

134. Doyle AC, Goldschmidt A, Huang C, et al. Reduction of overweight and eating disorder symptoms via the Internet in adolescents: a randomized controlled trial. *J Adolesc Health* 2008; 43: 172–179.

135. Haerens L, Deforche B, Maes L, et al. Body mass effects of a physical activity and healthy food intervention in middle schools. *Obesity (Silver Spring)* 2006; 14: 847–854.

136. Tate DF, Wing RR, Winett RA. Using Internet technology to deliver a behavioral weight loss program. *JAMA* 2001; 285: 1172–1177.

137. Marks JT, Campbell MK, Ward DS, et al. A comparison of web and print media for physical activity promotion among adolescent girls. *J Adolesc Health* 2006; 39: 96–104.

138. White MA, Martin PD, Newton RL, et al. Mediators of weight loss in a family-based intervention presented over the Internet. *Obes Res* 2004; 12: 1050–1059.

139. Williamson DA, Martin PD, White MA, et al. Efficacy of an Internet-based behavioral weight loss program for overweight adolescent African-American girls. *Eat Weight Disord* 2005; 10: 193–203.

140. Harvey-Berino J, Pintauro SJ, Gold EC. The feasibility of using Internet support for the maintenance of weight loss. *Behav Modif* 2002; 26: 103–116.

141. Harvey-Berino J, Pintauro S, Buzzell P, Gold EC. Effect of Internet support on the long-term maintenance of weight loss. *Obes Res* 2004; 12: 320–329.

142. Harvey-Berino J, Pintauro S, Buzzell P, et al. Does using the Internet facilitate the maintenance of weight loss? *Int J Obes* 2002; 26: 1254–1260.

143. Womble LG, Wadden TA, McGuckin BG, et al. A randomized controlled trial of a commercial Internet weight loss program. *Obes Res* 2004; 12: 1011–1018.

144. McConnon A, Kirk SF, Cockroft JE, et al. The Internet for weight control in an obese sample: results of a randomised controlled trial. *BMC Health Serv Res* 2007; 7: 206.

145. Webber KH, Tate DF, Bowling JM. A randomized comparison of two motivationally enhanced Internet behavioral weight loss programs. *Behav Res Ther* 2008; 46: 1090–1095.

146. Morgan PJ, Lubans DR, Collins CE, et al. The SHED-IT randomized controlled trial: evaluation of an Internet-based weight-loss program for men. *Obesity* 2009; 17: 2025–2032.

147. Krukowski RA, Harvey-Berino J, Ashikaga T, et al. Internet-based weight control: the relationship

between Web features and weight loss. *Telemed J E Health* 2008; 14: 775–782.

148. Norman GJ, Zabinski MF, Adams MA, et al. A review of eHealth interventions for physical activity and dietary behavior change. *Am J Prev Med* 2007; 33: 336–345.
149. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: a meta-analysis of 75 randomized controlled trials, 1988–2007. *Prev Med* 2008; 47: 3–16.
150. Saperstein SL, Atkinson NL, Gold RS. The impact of Internet use for weight loss. *Obes Rev* 2007; 8: 459–465.
151. Petersen R, Sill S, Lu C, et al. Effectiveness of employee Internet-based weight management program. *J Occup Environ Med* 2008; 50: 163–171.
152. Bish CL, Blanck HM, Serdula MK, et al. Diet and physical activity behaviors among Americans trying to lose weight: 2000 Behavioral Risk Factor Surveillance System. *Obes Res* 2005; 13: 596–607.

Dietary Approaches to Obesity

Gal Dubnov-Raz, MD, MSc,[1] and Elliot M. Berry MD, FRCP[2]

[1]Sheba Medical Center, Tel Hashomer, Israel
[2]Hebrew University-Hadassah Medical School, Jerusalem, Israel

OUTLINE

Diet and or Exercise?
What is the "Best Diet"?
  Low-Fat Versus Low-Carbohydrate Diets
  Low-Glycemic-Index Diets
  Calcium and Dairy Products
The Diets of Successful Weight-Losers
Conclusion

ABSTRACT

The treatment of obesity demands a negative energy balance, which obviously necessitates reducing the *quantity* of food intake. Yet the optimal diet composition *quality* for weight loss is far from known. The major controversy seems to be between choosing low-fat or low-carbohydrate diets, with additional debates on the importance of the glycemic index and the amounts of dairy products and calcium. It is evident from many large, randomized trials that the "best diet" that can suit everyone has yet to be identified. The different options for low-calorie diets should be prescribed on an individual basis, in order to best fit each patient's personal preferences. All hypocaloric regimens are suitable, as long as they work, for any given subject. The problem, of course, is long-term maintenance and adherence to an altered lifestyle. ***Mt Sinai J Med 77:488–498, 2010.*** **©** *2010 Mount Sinai School of Medicine*

Key Words: adiposity, carbohydrates, diet, fat, nutrition, obesity, overweight, weight loss.

> "Now to perform a true physician's part,
> And show I am perfect master of my art,
> I will prescribe what diet you should use,
> What food you ought to take, and what refuse."
> –Publius Ovidius Naso ("Ovid," 43 B.C.–A.D.17),
> *Remedia Amoris*

Address Correspondence to:

Gal Dubnov-Raz
The Edmond and Lily Safra
Children's Hospital
Sheba Medical Center
Tel Hashomer, Israel
Email: gal.dubnov-raz@sheba.
health.gov.il

The relentless, increasing prevalence of overweight and obesity demands an ongoing search for preventive and treatment strategies. Though it may seem to be only a recent issue, the struggle with excess weight and fat has troubled mankind for centuries. Galen, in the second century, already prescribed exercise and foods containing "little nourishment" when treating obesity.[1] About 800 years ago, Maimonides asserted that one should never eat unless hungry, that the stomach should only be filled to three-quarters full, and that "whenever one exercises and works–he will not become sick, and his strength will increase." He also recommended eating only once a day.[2] Although obesity existed in former times, it is only during the past decades that its dramatic increase in prevalence has reached pandemic levels.[3–7] Because genes do not change so rapidly, this phenomenon is believed to reflect an obesogenic or toxic environment, encouraging a continual positive energy balance. The rates of obesity-associated diseases have paralleled the increased prevalence of obesity.[8] This has occurred despite numerous studies, national statements, and position stands of professional organizations that obesity must be addressed seriously, prevented, and treated.[9] It seems that more than ever, multilevel, multidisciplinary strategies for obesity prevention and treatment must be used, with combined forces of governments, healthcare services, medical societies, legislation, industry, and media–but ultimately, it demands the efforts of the "end consumer": the obese individual.

More than ever, multilevel, multidisciplinary strategies for obesity prevention and treatment must be used, with combined forces of governments, healthcare services, medical societies, legislation, industry, and media–but ultimately, it demands the efforts of the "end consumer": the obese individual.

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20210

© 2010 Mount Sinai School of Medicine

The treatment of obesity demands a negative energy balance, achieved by reducing food intake in parallel with increasing energy expenditure. Even though this sounds so simple in theory and so obvious in thermodynamic terms, the decades of advocating weight loss have failed to stop the obesity epidemic. Though consuming less energy is a clear recommendation, there still remains the ongoing search for the proper nutrient combinations for treating obesity.

This review will discuss contemporary issues in the dietary treatment of obesity, based mainly on randomized trials of diets of varying compositions. Due to the numerous effects of selected nutrients and diets on the medical complications of obesity (eg, diabetes, hypertension, hypercholesterolemia, cardiovascular disease, cancer, and others), we will confine the discussion to weight loss per se in adults.

DIET AND/OR EXERCISE?

Weight loss will occur when a negative energy balance is obtained. This can be done by either eating less, exercising more, or combining both. A deficiency of 500 kcal/day is recommended in order to achieve a weight-loss rate of ~1 pound (~0.5 kg) per week.[9–11] It seems that the same energy deficit obtained by either increasing exercise or decreasing dietary intake yields similar changes in body weight.

> *The same energy deficit obtained by either increasing exercise or decreasing dietary intake yields similar changes in body weight.*

Several randomized trials have examined the differences between these 2 methods for weight loss.[14–17] Participants were divided into isocaloric groups, achieved by either decreased energy intake, increased energy expenditure, or combinations of both. Weight loss was similar in all groups, suggesting that when body weight is the outcome of interest, "a calorie is a calorie," whether burned by increased exercise or never consumed at all. The specific tissue composition of weight loss can reach 85% fat and 15% lean body mass (muscle and individual organs), whereas loss of the latter component causes a decrease in the resting metabolic rate.[18] Exercise aids in minimizing the loss of lean body mass, thereby reducing the decrease in metabolic rate that accompanies any loss in body weight.[12,13]

It must be remembered, however, that physical activity exerts numerous health benefits over and above affecting weight loss. It improves the metabolic complications of obesity, even without weight loss,[14,15] at the same time as it also prevents numerous chronic diseases such as heart disease,[19] cancer,[20] obesity,[21] and diabetes,[22] as well as death,[23] and improves mood and well-being.[24] In addition, large amounts of physical activity are essential for maintaining the reduced obese state,[25] so exercise must be an integral part of any weight-loss program, regardless of the diet regimen chosen.

WHAT IS THE "BEST DIET"?

For many years, researchers and clinicians have debated what should be the optimal diet for weight loss, and whether such a diet actually exists. Some key questions are whether a calorie ingested is indeed just a calorie, regardless of its source, and whether different compositions of isocaloric diets have different effects on weight loss.[26] The major controversy seems to be between low-fat (LF) and low-carbohydrate (LC) diets.[27] Additional contemporary issues concern the benefits of using diets with a low glycemic index (GI), or ones rich in dairy products and calcium.

Low-Fat Versus Low-Carbohydrate Diets

LF diets are those that restrict fat intake to 25%–35% of daily energy intake.[11,28,29] This is actually the recommended intake of fat for the general population. LC diets are those that advocate minimizing carbohydrate intake (~20–100 g per day, depending on the type of diet and its stage), with the additional calories needed originating from protein and fat. An overview of the constituents of common diets is available.[30]

Hypothesized Mechanisms for the Benefit of Low-Carbohydrate Diets

There are several suggested mechanisms by which LC diets could assist weight loss. These include the higher amount of protein, which promotes satiety more than carbohydrates; ongoing gluconeogenesis to compensate for the body's carbohydrate needs, which is an energy-consuming process; increased diuresis; loss of glycogen stores and their associated water; high levels of circulating ketones, which suppress appetite; narrowed food choices and thus a decreased energy intake; or merely underreporting of the food ingested, which causes a false belief that such a diet leads to less food consumption.[31–35]

Studies Comparing Low-Fat and Low-Carbohydrate Diets

Numerous studies have been conducted comparing these opposing approaches for weight loss, generating many published editorial comments, reviews, and meta-analyses.[27,36–56] Systematic reviews and meta-analyses reveal that in the long term (≥1 year) LC diets do not induce more weight loss than LF diets, with conflicting results regarding blood pressure, lipid profile, or glucose/insulin changes; attrition rates are usually similar.

> *Systematic reviews and meta-analyses reveal that in the long term (≥1 year) low-carbohydrate diets do not induce more weight loss than low-fat diets.*

A more recent systematic review showed that at 1 year, LC diets resulted in a mean weight loss of −1 kg more than LF diets.[56] Even a plant-based LC diet with no meat products did not promote more weight loss than an LF diet.[54] It should also be noted that after an LC intervention was stopped, long-term follow-up (2 years) did not identify a difference in body weight.[51] Moreover, participants in the LC group, who lost more weight during the study intervention, tended to regain more weight.

Several large, long-term, randomized trials comparing various diet regimens for weight loss, including LC and LF diets, have recently been performed (Table 1). In the study by Dansinger *et al.*,[47] 160 overweight participants were randomly assigned to 1 of 4 diets for 1 year: Atkins (carbohydrate restriction), Zone (macronutrient balance), Weight Watchers (calorie restriction), or Ornish (fat restriction). All diets resulted in modest weight loss at 1 year, with no statistically significant differences between diets, and no meaningful differences in heart disease risk factors or insulin resistance. In the study by Truby *et al.*,[48] 293 overweight participants were randomized to 1 of 5 diet groups for a 6-month trial: Atkins, the Slim-Fast plan, Weight Watchers, a precooked meal program (Rosemary Conley's), and a control group. All diets resulted in significant but similar loss of body fat and weight at the end, despite lack of control of total energy intake in the research protocols. Gardner *et al.*[49] randomized 311 overweight women to follow one of these diets: Atkins, Zone, Ornish, or LEARN (Lifestyle, Exercise, Attitudes, Relationships, and Nutrition), an LF diet based on national guidelines. At 12 months,

Table 1. *Summary of Long-Term (≥12 Months) Randomized Trials Comparing Low-Fat and Low-Carbohydrate Diets.*

			Low-Fat Group*†				Low-Carbohydrate Group†				Differences at Study End†	
Study	Sex	Duration (mo)	No. of Patients	Designated Fat (% en)	Target Maintained at Study End?	Mean Weight Loss (kg)	No. of Patients	Designated Carbohydrate Amount	Target Maintained at Study End?	Mean Weight Loss (kg)	Energy Intake	Weight Loss
Foster *et al.*[40]	M/F	12	30	25%	NA	−4.5	33	20 g/day initially‡	NA	−7.3	NA	No
Stern *et al.*[38]	M/F	12	43	30%	No	−3.1	44	30 g/day	No	−5.1	No	No
Dansinger *et al.*[47]	M/F	12	40	10%	No	−3.3	40	20–50 g/day	No	−2.1	No	No
Gardner *et al.*[49]	F	12	76	10%	No	−4.7	77	20–50 g/day	No	−2.6	No	No
Shai *et al.*[50]	M/F	24	104	30%	Yes	−3.3	109	20–120 g/day	NA	−5.5	No	Yes
Sacks *et al.*[55]	M/F	24	204	20%	Yes (26% en)	−4.1	201	35% en	No	−3.9	No	No
Brinkworth *et al.*[52]	M/F	12	36	30%	Yes	−11.5	33	20–40 g/day	Yes	−14.5	No	No

Abbreviations: % en, percent of total energy intake; F, female; M, male; NA, data not available.
* We chose the lowest-fat diet when several were available.
†We chose data on study completers only when available.
‡Individual increases allowed afterward.

weight loss was greater in the Atkins diet group compared with the other diet groups (−4.7 kg, −1.6 kg, −2.2 kg, and −2.6 kg, respectively) but significantly different only from the Zone diet. In the study by Sacks *et al.*,[55] 811 overweight adults were randomized to 1 of 4 diets for 2 years: a LC/high-protein diet, a LC/low-protein diet, a LF/high-protein diet, and a LF/low-protein diet. At 2 years, all diets produced a similar, clinically meaningful weight loss, regardless of the macronutrient ratios. Finally, a workplace-based study conducted in Israel among 322 overweight and obese employees compared the effect of either a LC, LF, or Mediterranean-style diet on weight loss and cardiovascular risk factors.[50] After 2 years, in the participants who adhered to the program, both the LC and the Mediterranean diet produced a larger weight loss compared with the LF diet (−0.5 body mass index [BMI] units in the LC and Mediterranean diets versus the LF diet). As these studies differ greatly in their settings, populations, dietary regimens, patient contact, and adherence, one should consider either recent meta-analyses or a specific study that is most similar to the relevant clinical setting regarding the optimal macronutrient composition for weight loss. An important issue indicated from the results in Table 1 is that in most long-term studies, even the participants who stayed in the study did not fully adhere to their designated diet composition. This has 2 implications. One is that a diet that demands large changes in eating habits, simply given randomly to patients, should not be expected to last for long periods of time. Secondly, we are still lacking long-term studies with low dropout rates, in which the participants adhered to the specific regimen throughout the study period. Currently, reasonable conclusions from all these studies might be that, when seeking weight loss, "a calorie is a calorie is a calorie" (with respect to Gertrude Stein), or that "any diet is efficacious–all the while it works. . . ."

A diet that demands large changes in eating habits, simply given randomly to patients, should not be expected to last for long periods of time.

The Safety of Low-Carbohydrate Diets

An issue for concern is the safety of long-term LC diets. A systematic review of >100 studies identified no adverse effects or unwarranted metabolic changes, especially in renal function.[45] Yet cases of severe, even life-threatening ketoacidosis in subjects following an Atkins-type diet have been described.[57,58] As most studies were conducted among healthy volunteers, and only a few among diabetic patients, it should be remembered that insufficient data exist regarding LC diets in populations with underlying diseases. A recent study in mice has shown that a high-protein, LC diet might promote growth of atherosclerotic plaques, independent of the classical known risk factors. These safety issues seem relevant to man[59]; therefore, apart from the benefits of weight loss in following LC diets, the impact on cardiovascular disease should also be monitored.

Apart from the benefits of weight loss in following low-carbohydrate diets, the impact on cardiovascular disease should also be monitored.

In summary, despite several hypothetical benefits of LC diets over LF diets, even high-quality, rigorous studies continue to generate conflicting results. Future research will probably assist in identifying the optimal proportions between macronutrients for weight loss and metabolic improvements. Nevertheless, the most important factor determining the success of any diet is the ability to maintain long-term adherence to the new lifestyle changes involved. This is an individual choice, affected by numerous factors, both extrinsic and intrinsic to the person. We suggest that the overweight individual seeking a weight-loss diet be informed of the various options of macronutrient proportions and choose that which fits him best.

The most important factor determining the success of any diet is the ability to maintain long-term adherence to the new lifestyle changes involved.

Low-Glycemic-Index Diets

The GI is a measure of glucose availability of different foods,[60] and is calculated as the percentage of the area under the plasma glucose response curve measured during the 2 hours after the ingestion of the food in question, relative to the same amount of carbohydrate taken as glucose/white bread. In foods with a high GI (HGI; >70), such as white bread, potatoes, or corn flakes, glucose is readily digested and absorbed, causing a rapid increase in plasma glucose levels and, correspondingly, a rapid increase in insulin levels. Foods with a low GI (LGI; <50), such as pasta, fruits, and beans, cause a much slower increment in plasma glucose levels.

Hypothesized Mechanism of Low-Glycemic-Index Diets

It has been postulated that the reason for the failure of fat reduction in combating the current obesity epidemic lies in the carbohydrate content. The substitution of fat with the highly prevalent HGI foods merely causes increased hunger, anabolism in adipose tissue, and weight gain.[61]

It has been postulated that the reason for the failure of fat reduction in combating the current obesity epidemic lies in the carbohydrate content. The substitution of fat with the highly prevalent high-glycemic-index foods merely causes increased hunger, anabolism in adipose tissue, and weight gain.

The rapid absorption of sugar from a HGI diet causes a rapid surge in insulin secretion, which then exerts its anabolic effects. In addition, the high insulin levels decrease blood glucose levels, causing more hunger only a few hours after a HGI meal. Following a LGI diet, insulin secretion is lower, with lower postprandial hunger.[61]

Studies Comparing Low- and High-Glycemic-Index Diets

Several studies investigated whether a LGI diet is superior to a HGI diet for weight loss. A prospective cohort study of 6-year follow-up was conducted among 185 men and 191 women, examining the role of GI in future weight and fat-mass changes.[62] Using questionnaires for food frequency and various confounders, the authors concluded that, only in women, a HGI diet was indeed associated with increases in BMI and fat mass. Another observational study among 572 healthy adults also showed that changes in BMI were positively associated with dietary intake of carbohydrates and the GI.[63]

Interventional studies have also been conducted to examine whether lowering the GI can promote weight loss. A meta-analysis is available of the highest-quality randomized studies that examined the effect of LGI diets on weight loss.[64] These diets produce larger decreases in BMI, fat mass, and total and low-density lipoprotein cholesterol. As studies were generally short-term, further research with longer follow-up was recommended, and later randomized trials did not reveal an added benefit of such LGI diets. Very few long-term studies comparing LGI with HGI diets exist. In an 18-month study from Brazil, 203 healthy women with a BMI of 23–30 kg/m^2 were randomized to either a LGI or a HGI diet,[65] both with high carbohydrate content (60% of energy), low fat (26%–28% of energy), and with a small energy deficit (−100–300 kcal/day). The LGI group had a slightly greater weight loss in the first 2 months of follow-up, but both groups began to regain weight after 12 months, and at 18 months only a small weight loss was maintained (−260 g and −400 g for LGI and HGI diets, respectively), which did not differ significantly between groups. Both groups maintained their designated GI value, which averaged 30 in the LGI group and 72 in the HGI group. Another recent, shorter-term study examined the effect of a LGI diet on weight-loss maintenance.[66] After a 6% weight-loss phase, 42 participants were randomized to a 4-month weight-maintenance diet with either a LGI or HGI. There were no statistically significant differences. In another study, researchers randomized 34 healthy overweight adults to diets with 30% caloric deficits and either a HGI or a LGI.[67] All food was provided for 6 months, and subjects self-administered the diet plans for 6 additional months. The LGI diet had 40% daily energy as carbohydrates (compared with 60% in the HGI) and a 33-unit-lower GI value. Body weight and fatness did not differ between diet groups at any point during the study. Additional outcomes included energy intake, hunger, satiety, and resting metabolic rate–none of which differed between the groups at the end of the study. Another study, conducted among 39 obese subjects randomized to hypocaloric diets (−760 kcal/day) with LGI, HGI, or high fat, concluded that the GI had no effect on weight loss.[68]

In summary, high-quality randomized trials continue to raise conflicting results regarding the role of the GI in weight loss and maintenance.

High-quality randomized trials continue to raise conflicting results regarding the role of the glycemic index in weight loss and maintenance.

Several methodological issues may explain these findings. One major concern is that HGI and LGI signify a range of glycemic indices, and not 1 single figure. Another issue is the use of food-frequency questionnaires, which can limit accuracy and insert many forms of bias. We suggest that more studies be conducted in free-living conditions, where the presumed higher satiety of a LGI diet can be translated into a greater reduction of energy intake; in a controlled trial, holding the energy intake constant in both LGI and HGI groups might blunt one of the possible benefits of a LGI diet. In addition, we suggest that in the clinic setting, the effect of eating LGI and HGI meals on hunger and satiety be examined on an individual basis, to see if lowering the GI may assist in the reduced-calorie diet.

Calcium and Dairy Products

In 1984, McCarron reported an inverse relationship between dietary calcium intake and body weight.[69] Since then, several observational studies and interventions have assessed the reproducibility, magnitude, and mechanisms of this phenomenon.

Hypothesized Mechanism of Calcium in Weight Control

Several possible mechanisms for the effect of dietary calcium and dairy foods on cell metabolism and adiposity have been identified.[70,71] The major one seems to be a suppression of intracellular calcium, which then enhances lipolysis, with an interesting greater effect with dairy products than calcium supplements. Increased fecal fat loss is another mechanism by which calcium may aid in weight loss.[70]

Studies Comparing Diets of Varying Calcium Content

Studies conducted in this field show conflicting results regarding diets with a higher calcium or dairy intake, showing either a lower weight (in cross-sectional studies) or increased weight loss (in longitudinal studies),[72–81] no effect,[81–87] or even increased weight or weight gain.[86,88,89] A systematic review of 13 randomized trials involving supplementation of either calcium or dairy products found conflicting results, and summarized that such diets have no effect on weight loss.[90] A recent randomized, placebo-controlled trial in 340 overweight and obese participants examined the role of calcium 1500 mg/day for 2 years on weight loss.[87] No statistically significant differences between the calcium and placebo groups in change in body weight, BMI, or body fat mass were found.

A systematic review of 13 randomized trials involving supplementation of either calcium or dairy products found conflicting results, and summarized that such diets have no effect on weight loss.

An important pitfall when recommending eating more dairy products for weight loss might be an increased caloric intake. In one study, consuming 3 dairy products per day averaged an additional 200 kcal/day[91]—a meaningful increase when considered for long-term management. Data from the Health Professionals Follow-Up Study showed that men with the largest increase in high-fat dairy products gained significantly more weight than did men who decreased their intake, showing that total energy consumption must be taken into consideration to prevent its counteracting of possible beneficial effects of dairy products on weight.[88] In addition, various dairy products may affect body weight differently, thereby suggesting that pooling all "dairy products" in questionnaires or diet recommendations may be erroneous. In a large study of 19,352 women,[92] researchers examined the association between changes in dairy-product consumption and weight change over 9 years while addressing differential intakes of whole milk (3% fat), medium-fat milk (1.5% fat), low-fat milk (≤0.5% fat), whole sour milk (3% fat), low-fat sour milk (0.5% fat), and cheese and butter (80% fat). Findings were that increasing intake of dairy products and calcium did not universally affect weight gain, whether these increases involved low-fat milk or cheese and butter. A significantly lower risk for gaining >1 kg/year was observed in subjects who consumed >1 portion of 3% milk (relative risk, 0.85) and who maintained a high cheese intake (relative risk, 0.70). These results were not affected by correction for calcium or energy intake. Therefore, it may be concluded that not all milk products are created equal, and future studies may identify the more beneficial dairy products (if any) in assisting weight loss.

THE DIETS OF SUCCESSFUL WEIGHT-LOSERS

We are all aware of successful weight-losers; how did they do it? The National Weight Control Registry

(NWCR) can assist in identifying the techniques used by people who have managed to maintain their weight loss.[93] Fifty-five percent of registry members lost weight through formal programs and 45% lost weight on their own, suggesting that an official program is not mandatory for long-term weight loss. Eighty-nine percent used both diet and exercise for weight loss, 10% used only diet, and 1% used only exercise. The 3 most widely used methods for limiting food intake were to limit intake of certain types or classes of foods (88%), to limit quantities of food eaten (44%), and to count calories (44%). More than 40% of the participants reported that maintaining their weight loss was less difficult than losing weight initially, which gives hope to all those commencing a weight-loss program.

We are all aware of successful weight-losers; how did they do it? According to the National Weight Control Registry, 89% used both diet and exercise for weight loss, 10% used only diet, and 1% used only exercise. The 3 most widely used methods for limiting food intake were to limit intake of certain types or classes of foods (88%), to limit quantities of food eaten (44%), and to count calories (44%).

The average diet composition of the NWCR members may shed light on what works best on a population level, in a real-life setting. In a report from 1997,[93] registry members consumed 24% of their energy from fat, 19% from protein, and 56% from carbohydrates. Examining their diet composition in 2003[94] showed that registry members consumed more energy from fat (29%) and less from carbohydrates (49%). The proportional intake of protein remained unchanged, at 19%. Mean BMIs did not differ between these 2 time points, yet subjects were heavier and lost more weight before registry enrollment.[94] This change in diet composition could be related to methodological issues, but it could also reflect a true growing interest in LC diets. In a study from 2007, only 11% of NWCR members reported using a LC diet,[95] but their very high fat intake was enough to markedly raise the mean of the whole cohort. Comparing their lifestyle habits, LC diet users reported consuming more total energy, exhibited less dietary restraint, and were much less physically active than other registry members. No differences in 3-year weight regain were observed between LC diet users and other NWCR members, suggesting that consuming such a diet does not assist in weight-loss maintenance better than other diet forms. Most members of the NWCR admitted to doing a large amount of exercise each day,[25] which attests to their rigorous change in lifestyle.

No differences in 3-year weight regain were observed between National Weight Control Registry members who used a low-carbohydrate diet and other registry members, suggesting that consuming such a diet does not assist in weight-loss maintenance better than other diet forms. Most members of the registry admitted to doing a large amount of exercise each day, which attests to their rigorous change in lifestyle.

CONCLUSION

In no other medical therapeutic area is the theory so simple as in weight control. Yet the major gap lies in translating this theory into practice. A key message from this review is that randomizing study participants to consume meticulously constructed diets works in very few cases. We believe that the answer lies in personal choice, as many diets seem to work. It is obvious that a reduced-calorie diet is essential–but its macronutrient composition, GI, and calcium content remain to be individually tailored. Further, whichever diet is chosen, it must be accompanied by an exercise program for long-term maintenance of the reduced weight.[25]

What determines individual diet preference and effectiveness? These could be biological or genetic factors. For example, the rate of insulin secretion has been found to modify the effect of a different glycemic load[96] or different fat intakes[97] on weight loss. Perhaps measuring insulin secretion following standardized test meals could predict responders to a specific diet regimen. The numerous possible diet-gene interactions are now becoming clearer. For example, carriers of the A allele of the *FTO* gene, who have increased body fat and weight, responded better to a 3-year Mediterranean diet than did wild-type subjects.[98] Such studies suggest that in the near future, genotyping obese subjects may have clinical implications and will allow tailoring of more effective diet compositions for weight loss.

In conclusion, until the "best diet" for weight loss is identified, the different options of reduced-calorie diets should probably be personalized to the abilities and preferences of the individual. Additional efforts should be made to further enhance compliance with the diet chosen by promoting social, group, and spouse support; behavior modifications; psychotherapy; weekly calls, text messaging, or e-mails; or reinforcement meetings with a nutritionist or personal trainer. Governmental actions on many levels are also needed in order to counter the temptations inherent to our modern obesogenic environment. These may take the form of legislation against advertising to children, providing safe walkways and bicycle paths, and, above all, giving obesity prevention a very high priority on the national health agenda.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Guggenheim KY. *Basic Issues of the History of Nutrition*. Jerusalem, Israel: Akademia University Press; 1990.
2. Muntner S. *The Medical Writings of Moses Maimonides: Treatise on Asthma*. Philadelphia, PA: J.B. Lippincott Co.; 1963; 28.
3. Li C, Ford ES, McGuire LC, et al. Increasing trends in waist circumference and abdominal obesity among US adults. *Obesity* 2007; 15: 216–224.
4. Wardle J, Boniface D. Changes in the distributions of body mass index and waist circumference in English adults, 1993/1994 to 2002/2003. *Int J Obes (Lond)* 2008; 32: 527–532.
5. Berg C, Rosengren A, Aires N, et al. Trends in overweight and obesity from 1985 to 2002 in Goteborg, West Sweden. *Int J Obes (Lond)* 2005; 29: 916–924.
6. Bendixen H, Holst C, Sorensen TI, et al. Major increase in prevalence of overweight and obesity between 1987 and 2001 among Danish adults. *Obes Res* 2004; 12: 1464–1472.
7. Koziel S, Szklarska A, Bielicki T, et al. Changes in the BMI of Polish conscripts between 1965 and 2001: secular and socio-occupational variation. *Int J Obes (Lond)* 2006; 30: 1382–1388.
8. Ford ES, Giles WH, Mokdad AH. Increasing prevalence of the metabolic syndrome among US adults. *Diabetes Care* 2004; 27: 2444–2449.
9. National Institutes of Health. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: the Evidence Report [published correction appears in *Obes Res* 1998; 6: 464]. *Obes Res* 1998; (6 suppl 2): S51–S210.
10. Avenell A, Sattar N, Lean M. ABC of obesity. Management: Part I—behaviour change, diet, and activity. *BMJ* 2006; 333: 740–743.
11. Klein S, Sheard NF, Pi-Sunyer X, et al. Weight management through lifestyle modification for the prevention and management of type 2 diabetes: rationale and strategies. A statement of the American Diabetes Association, the North American Association for the Study of Obesity, and the American Society for Clinical Nutrition. *Am J Clin Nutr* 2004; 80: 257–263.
12. Leibel RL, Hirsch J. Diminished energy requirements in reduced-obese patients. *Metabolism* 1984; 33: 164–170.
13. Leibel RL, Berry EM, Hirsch J. Metabolic and hemodynamic responses to endogenous and exogenous catecholamines in formerly obese subjects. *Am J Physiol* 1991; 260: R785–R791.
14. Ross R, Dagnone D, Jones PJ, et al. Reduction in obesity and related comorbid conditions after diet-induced weight loss or exercise-induced weight loss in men: a randomized, controlled trial. *Ann Intern Med* 2000; 133: 92–103.
15. Ross R, Janssen I, Dawson J, et al. Exercise-induced reduction in obesity and insulin resistance in women: a randomized controlled trial. *Obes Res* 2004; 12: 789–798.
16. Redman LM, Heilbronn LK, Martin CK, et al. Effect of calorie restriction with or without exercise on body composition and fat distribution. *J Clin Endocrinol Metab* 2007; 92: 865–872.
17. Fontana L, Villareal DT, Weiss EP, et al. Calorie restriction or exercise: effects on coronary heart disease risk factors. A randomized, controlled trial. *Am J Physiol Endocrinol Metab* 2007; 293: E197–E202.
18. Bosy-Westphal A, Kossel E, Goele K, et al. Contribution of individual organ mass loss to weight loss–associated decline in resting energy expenditure. *Am J Clin Nutr* 2009; 90: 993–1001.
19. Haskell WL, Lee IM, Pate RR, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation* 2007; 116: 1081–1093.
20. Kushi LH, Byers T, Doyle C, et al. American Cancer Society Guidelines on Nutrition and Physical Activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity [published correction appears in *CA Cancer J Clin* 2007;57:66]. *CA Cancer J Clin* 2006; 56: 254–281.
21. Saris WH, Blair SN, van Baak MA, et al. How much physical activity is enough to prevent unhealthy weight gain? Outcome of the IASO 1st Stock Conference and consensus statement. *Obes Rev* 2003; 4: 101–114.
22. Sigal RJ, Kenny GP, Wasserman DH, et al. Physical activity/exercise and type 2 diabetes: a consensus statement from the American Diabetes Association. *Diabetes Care* 2006; 29: 1433–1438.
23. Lee CD, Blair SN, Jackson AS. Cardiorespiratory fitness, body composition, and all-cause and cardiovascular disease mortality in men. *Am J Clin Nutr* 1999; 69: 373–380.

24. Dubnov G, Berry EM. Physical activity and mood: the endocrine connection. In: Warren MP, Constantini N, eds. *Sports Endocrinology*. NJ: Humana Press; 2000; 421–432.
25. Catenacci VA, Ogden LG, Stuht J, et al. Physical activity patterns in the National Weight Control Registry. *Obesity* 2008; 16: 153–161.
26. Dubnov-Raz G, Berry EM. The dietary treatment of obesity. *Endocrinol Metab Clin North Am* 2008; 37: 873–886.
27. Foreyt JP, Salas-Salvado J, Caballero B, et al. Weight-reducing diets: are there any differences? *Nutr Rev* 2009; 67: (suppl 1): S99–S101.
28. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation* 2002; 106: 3143–3421.
29. American Heart Association Nutrition Committee. Diet and lifestyle recommendations revision 2006: a scientific statement from the American Heart Association Nutrition Committee. *Circulation* 2006; 114: 82–96.
30. Freedman MR, King J, Kennedy E. Popular diets: a scientific review. *Obes Res* 2001; 9: (suppl 1): 1S–40S.
31. Astrup A, Meinert Larsen T, Harper A. Atkins and other low-carbohydrate diets: hoax or an effective tool for weight loss? *Lancet* 2004; 364: 897–899.
32. Feinman RD, Fine EJ. "A calorie is a calorie" violates the second law of thermodynamics. *Nutr J* 2004; 3: 9.
33. Buchholz AC, Schoeller DA. Is a calorie a calorie? *Am J Clin Nutr* 2004; 79: S899–S906.
34. Schoeller DA, Buchholz AC. Energetics of obesity and weight control: does diet composition matter? *J Am Diet Assoc* 2005; 105: (suppl 1): S24–S28.
35. Brehm BJ, Spang SE, Lattin BL, et al. The role of energy expenditure in the differential weight loss in obese women on low-fat and low-carbohydrate diets. *J Clin Endocrinol Metab* 2005; 90: 1475–1482.
36. Brehm BJ, Seeley RJ. Daniels SR, D'Alessio DA. A randomized trial comparing a very low carbohydrate diet and a calorie-restricted low fat diet on body weight and cardiovascular risk factors in healthy women. *J Clin Endocrinol Metab* 2003; 88: 1617–1623.
37. McAuley KA, Hopkins CM, Smith KJ, et al. Comparison of high-fat and high-protein diets with a high-carbohydrate diet in insulin-resistant obese women. *Diabetologia* 2005; 48: 8–16.
38. Stern L, Iqbal N, Seshadri P, et al. The effects of low-carbohydrate versus conventional weight loss diets in severely obese adults: one-year follow-up of a randomized trial. *Ann Intern Med* 2004; 140: 778–785.
39. Yancy WS Jr, Olsen MK, Guyton JR, et al. A low-carbohydrate, ketogenic diet versus a low-fat diet to treat obesity and hyperlipidemia: a randomized, controlled trial. *Ann Intern Med* 2004; 140: 769–777.
40. Foster GD, Wyatt HR, Hill JO, et al. A randomized trial of a low-carbohydrate diet for obesity. *N Engl J Med* 2003; 348: 2082–2090.
41. Samaha FF, Iqbal N, Seshadri P, et al. A low-carbohydrate as compared with a low-fat diet in severe obesity. *N Engl J Med* 2003; 348: 2074–2081.
42. Segal-Isaacson CJ, Johnson S, Tomuta V, et al. A randomized trial comparing low-fat and low-carbohydrate diets matched for energy and protein. *Obes Res* 2004; 12: (suppl 2): 130S–140S.
43. Volek JS, Sharman MJ, Gomez AL, et al. Comparison of a very low-carbohydrate and low-fat diet on fasting lipids, LDL subclasses, insulin resistance, and postprandial lipemic responses in overweight women. *J Am Coll Nutr* 2004; 23: 177–184.
44. Pirozzo S, Summerbell C, Cameron C, et al. Advice on low-fat diets for obesity. *Cochrane Database Syst Rev* 2002; 2: CD003640.
45. Bravata DM, Sanders L, Huang J, et al. Efficacy and safety of low-carbohydrate diets: a systematic review. *JAMA* 2003; 289: 1837–1850.
46. Nordmann AJ, Nordmann A, Briel M, et al. Effects of low-carbohydrate vs low-fat diets on weight loss and cardiovascular risk factors: a meta-analysis of randomized controlled trials. *Arch Intern Med* 2006; 166: 285–293.
47. Dansinger ML, Gleason JA, Griffith JL, et al. Comparison of the Atkins, Ornish, Weight Watchers, and Zone diets for weight loss and heart disease risk reduction: a randomized trial. *JAMA* 2005; 293: 43–53.
48. Truby H, Baic S, deLooy A, et al. Randomised controlled trial of four commercial weight loss programmes in the UK: initial findings from the BBC "diet trials". *BMJ* 2006; 332: 1309–1314.
49. Gardner CD, Kiazand A, Alhassan S, et al. Comparison of the Atkins, Zone, Ornish, and LEARN diets for change in weight and related risk factors among overweight premenopausal women: the A TO Z Weight Loss Study: a randomized trial [published correction appears in *JAMA* 2007; 298: 178]. *JAMA* 2007; 297: 969–977.
50. Shai I, Schwarzfuchs D, Henkin Y, et al. Weight loss with a low-carbohydrate, Mediterranean, or low-fat diet. *N Engl J Med* 2008; 359: 229–241.
51. Vetter ML, Iqbal N, Dalton-Bakes C, et al. Long-term effects of low-carbohydrate versus low-fat diets in obese persons. *Ann Intern Med* 2010; 152: 334–335.
52. Brinkworth GD, Noakes M, Buckley JD, et al. Long-term effects of a very-low-carbohydrate weight loss diet compared with an isocaloric low-fat diet after 12 mo. *Am J Clin Nutr* 2009; 90: 23–32.
53. Delbridge EA, Prendergast LA, Pritchard JE, Proietto J. One-year weight maintenance after significant weight loss in healthy overweight and obese subjects: does diet composition matter? *Am J Clin Nutr* 2009; 90: 1203–1214.
54. Jenkins DJ, Wong JM, Kendall CW, et al. The effect of a plant-based low-carbohydrate ("Eco-Atkins") diet on body weight and blood lipid concentrations in hyperlipidemic subjects. *Arch Intern Med* 2009; 169: 1046–1054.
55. Sacks FM, Bray GA, Carey VJ, et al. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N Engl J Med* 2009; 360: 859–873.
56. Hession M, Rolland C, Kulkarni U, et al. Systematic review of randomized controlled trials of low-carbohydrate vs. low-fat/low-calorie diets in the management of obesity and its comorbidities. *Obes Rev* 2009; 10: 36–50.

57. Shah P, Isley WL. Ketoacidosis during a low-carbohydrate diet. *N Engl J Med* 2006; 354: 97–98.
58. Chen TY, Smith W, Rosenstock JL, et al. A life-threatening complication of Atkins diet. *Lancet* 2006; 367: 958.
59. Smith SR. A look at the low-carbohydrate diet. *N Engl J Med* 2009; 361: 2286–2288.
60. Jenkins DJ, Wolever TM, Taylor RH, et al. Glycemic index of foods: a physiological basis for carbohydrate exchange. *Am J Clin Nutr* 1981; 34: 362–366.
61. Brand-Miller JC, Holt SH, Pawlak DB, McMillan J. Glycemic index and obesity. *Am J Clin Nutr* 2002; 76: 281S–285S.
62. Hare-Bruun H, Flint A, Heitmann BL. Glycemic index and glycemic load in relation to changes in body weight, body fat distribution, and body composition in adult Danes. *Am J Clin Nutr* 2006; 84: 871–879.
63. Ma Y, Olendzki B, Chiriboga D, et al. Association between dietary carbohydrates and body weight. *Am J Epidemiol* 2005; 161: 359–367.
64. Thomas DE, Elliott EJ, Baur L. Low glycaemic index or low glycaemic load diets for overweight and obesity. *Cochrane Database Syst Rev* 2007; 3: CD005105.
65. Sichieri R, Moura AS, Genelhu V, et al. An 18-mo randomized trial of a low-glycemic-index diet and weight change in Brazilian women. *Am J Clin Nutr* 2007; 86: 707–713.
66. Philippou E, Neary NM, Chaudhri O, et al. The effect of dietary glycemic index on weight maintenance in overweight subjects: a pilot study. *Obesity* 2009; 17: 396–401.
67. Das SK, Gilhooly CH, Golden JK, et al. Long-term effects of 2 energy-restricted diets differing in glycemic load on dietary adherence, body composition, and metabolism in CALERIE: a 1-y randomized controlled trial. *Am J Clin Nutr* 2007; 85: 1023–1030.
68. Raatz SK, Torkelson CJ, Redmon JB, et al. Reduced glycemic index and glycemic load diets do not increase the effects of energy restriction on weight loss and insulin sensitivity in obese men and women. *J Nutr* 2005; 135: 2387–2391.
69. McCarron DA. Dietary calcium as an antihypertensive agent. *Nutr Rev* 1984; 42: 223–225.
70. Zemel MB. The role of dairy foods in weight management. *J Am Coll Nutr* 2005; 24: (suppl 6): S537–S546.
71. Teegarden D. The influence of dairy product consumption on body composition. *J Nutr* 2005; 135: 2749–2752.
72. Davies KM, Heaney RP, Recker RR, et al. Calcium intake and body weight. *J Clin Endocrinol Metab* 2000; 85: 4635–4638.
73. Lin YC, Lyle RM, McCabe LD, et al. Dairy calcium is related to changes in body composition during a two year exercise intervention in young women. *J Am Coll Nutr* 2000; 19: 754–760.
74. Zemel MB, Shi H, Greer B, et al. Regulation of adiposity by dietary calcium. *FASEB J* 2000; 14: 1132–1138.
75. Pereira MA, Jacobs DR, Van Horn L, et al. Dairy consumption, obesity, and the insulin resistance syndrome in young adults: the CARDIA study. *JAMA* 2002; 287: 2081–2089.
76. Mirmiran P, Ezmaillzadeh A, Azizi F. Dairy consumption and body mass index: an inverse relationship. *Int J Obes (Lond)* 2005; 29: 115–121.
77. Jacqmain M, Doucet E, Després JP, et al. Calcium intake, body composition, and lipoprotein-lipid concentrations in adults. *Am J Clin Nutr* 2003; 77: 1448–1152.
78. Shahar DR, Abel R, Elhayany A, et al. Does dairy calcium intake enhance weight loss among overweight diabetic patients? *Diabetes Care* 2007; 30: 485–489.
79. Lovejoy JC, Champagne CM, Smith SR, et al. Ethnic differences in dietary intakes, physical activity, and energy expenditure in middle-aged, premenopausal women: the Healthy Transitions Study. *Am J Clin Nutr* 2001; 74: 90–95.
80. Loos RJ, Rankinen T, Leon AS, et al. Calcium intake is associated with adiposity in black and white men and white women of the HERITAGE Family Study. *J Nutr* 2004; 134: 1772–1778.
81. Lorenzen JK, Molgaard C, Michaelsen KF, Astrup A. Calcium supplementation for 1 y does not reduce body weight or fat mass in young girls. *Am J Clin Nutr* 2006; 83: 18–23.
82. Phillips SM, Bandini LG, Cyr H, et al. Dairy food consumption and body weight and fatness studied longitudinally over the adolescent period. *Int J Obes Relat Metab Disord* 2003; 27: 1106–1113.
83. Venti CA, Tataranni PA, Salbe AD. Lack of relationship between calcium intake and body size in an obesity-prone population. *J Am Diet Assoc* 2005; 105: 1401–1407.
84. Gunther CW, Legowski PA, Lyle RM, et al. Dairy products do not lead to alterations in body weight or fat mass in young women in a 1-y intervention. *Am J Clin Nutr* 2005; 81: 751–756.
85. Snijder MB, van der Heijden AA, van Dam RM, et al. Is higher dairy consumption associated with lower body weight and fewer metabolic disturbances? The Hoorn Study. *Am J Clin Nutr* 2007; 85: 989–995.
86. Kamycheva E, Joakimsen RM, Jorde R. Intakes of calcium and vitamin D predict body mass index in the population of Northern Norway. *J Nutr* 2003; 133: 102–106.
87. Yanovski JA, Parikh SJ, Yanoff LB, et al. Effects of calcium supplementation on body weight and adiposity in overweight and obese adults: a randomized trial. *Ann Intern Med* 2009; 150: 821–829.
88. Rajpathak SN, Rimm EB, Rosner B, et al. Calcium and dairy intakes in relation to long-term weight gain in US men. *Am J Clin Nutr* 2006; 83: 559–566.
89. Berkey CS, Rockett HR, Willett WC, et al. Milk, dairy fat, dietary calcium, and weight gain: a longitudinal study of adolescents. *Arch Pediatr Adolesc Med* 2005; 159: 543–550.
90. Trowman R, Dumville JC, Hahn S, et al. A systematic review of the effects of calcium supplementation on body weight. *Br J Nutr* 2006; 95: 1033–1038.
91. Hollis JH, Mattes RD. Effect of increased dairy consumption on appetitive ratings and food intake. *Obesity (Silver Spring)* 2007; 15: 1520–1526.
92. Rosell M, Håkansson NN, Wolk A. Association between dairy food consumption and weight change over 9 y in 19,352 perimenopausal women. *Am J Clin Nutr* 2006; 84: 1481–1488.
93. Klem ML, Wing RR, McGuire MT, et al. A descriptive study of individuals successful at long-term

maintenance of substantial weight loss. *Am J Clin Nutr* 1997; 66: 239–246.

94. Phelan S, Wyatt HR, Hill JO, Wing RR. Are the eating and exercise habits of successful weight losers changing? *Obesity* 2006; 14: 710–716.
95. Phelan S, Wyatt H, Nassery S, et al. Three-year weight change in successful weight losers who lost weight on a low-carbohydrate diet. *Obesity* 2007; 15: 2470–2477.
96. Pittas AG, Das SK, Hajduk CL, et al. A low-glycemic load diet facilitates greater weight loss in overweight adults with high insulin secretion but not in overweight adults with low insulin secretion in the CALERIE Trial. *Diabetes Care* 2005; 28: 2939–2941.
97. Chaput JP, Tremblay A, Rimm EB, et al. A novel interaction between dietary composition and insulin secretion: effects on weight gain in the Quebec Family Study. *Am J Clin Nutr* 2008; 87: 303–309.
98. Razquin C, Martinez JA, Martinez-Gonzalez MA, et al. A 3-year intervention with a Mediterranean diet modified the association between the rs9939609 gene variant in FTO and body weight changes. *Int J Obes (Lond)* 2010; 34: 266–272.

A Personalized Approach to Metabolic Aspects of Obesity

Stephen A. Brietzke, MD, FACP

University of Missouri-Columbia, Columbia, MO

OUTLINE

Distinguishing "Benign" From "Toxic" Fat
The Metabolic Syndrome Concept
"Lean Healthy" Versus "Fat Toxic" Adipocytes: Metabolic Yin and Yang
Treatments to Prevent Type 2 Diabetes Mellitus
Lifestyle Modifications in the Treatment of Hypertension
Treatment of Dyslipidemia
Treatment of Hypertension and Prehypertension
Antihypertensive Choices in Hypertension and Prehypertension
Summary of Recommendations for Treatment of Patients with Features of Metabolic Syndrome
Conclusion

ABSTRACT

Metabolic syndrome, which is entwined in semantic controversy as to its actual existence as a distinct entity, links several important health conditions with obesity, and more specifically, excessive visceral adiposity. The most common linked disease states include type 2 diabetes mellitus, hypertension, dyslipidemia, obstructive sleep apnea, and cardiovascular and coronary heart disease. Much of the controversy surrounding the metabolic syndrome case definition is the purported centrality of insulin resistance as root cause, there being no universally agreed-upon standard for measurement of insulin resistance. Over the past decade, the visceral adipocyte itself has emerged as a key contributor rather than passive bystander in the genesis of the metabolic syndrome. Rather than being a simple storage bin for excess triglyceride, the visceral adipocyte is an active endocrine cell secreting a variety of signal hormones known in the aggregate as adipokines. In optimal health, the predominant recognized adipokine is adiponectin, with downstream insulin-sensitizing, anti-inflammatory, antithrombotic, provasodilatory effects systemically. By contrast, metabolic syndrome is characterized by reduced adiponectin and increased inflammatory adipokine secretion, with downstream effects of insulin resistance, heightened inflammation, prothrombosis, and vasoconstriction. These alternative metabolic states of the adipocyte are characterized in this review as metabolic "yin" and "yang." Lifestyle modifications and drug therapies that promote weight loss, increased physical exercise activity, and increased adiponectin production tend to modulate the system favorably toward metabolic "yin." ***Mt Sinai J Med 77:499–510, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: adiponectin, cardiovascular risk, dyslipidemia, insulin resistance, metabolic syndrome, visceral obesity.

Address Correspondence to:

Stephen A. Brietzke
Division of Endocrinology, Diabetes, and Metabolism
Department of Internal Medicine
University of Missouri-Columbia
Columbia, MO
Email: brietzkes@health.missouri.edu

Currently, two-thirds of the US population is either overweight or obese, which is important because of a constellation of diseases that are either caused or exacerbated by obesity. When named separately, these include hypertension, type 2 diabetes mellitus (T2DM), dyslipidemia, obstructive sleep apnea, and cardiovascular disease (CVD). When several of these

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20206

© 2010 Mount Sinai School of Medicine

conditions are present in the same individual (as is frequently the case), the term "metabolic syndrome" is often used, implying the existence of a thematic root cause of the component diseases. The most outward obvious root cause would be obesity itself. But not all obese patients have metabolic syndrome, and some patients who do not appear to be obese actually do exhibit features of the metabolic syndrome. A key feature of metabolic syndrome appears to be conversion of benign adiposity to a syndrome of "toxic fat." The purpose of this review is first to summarize emerging pathophysiologic concepts that directly link the biology of adipose (fat) tissue to health and disease, and then to examine evidence for treatments that improve the derangements in metabolically "toxic" obesity, ie, the metabolic syndrome. An individualized risk appraisal is important, because equal body fat does not necessarily translate into equal risk of disease expression.

DISTINGUISHING "BENIGN" FROM "TOXIC" FAT

To foster an understanding of unhealthy types of obesity, the physician might discuss with the patient fundamental differences between 2 distinct phenotypes of patients with equal degrees of obesity (for example, 2 patients with a body mass index [BMI] of 31 kg/m^2). One patient has symmetrical distribution of subcutaneous adiposity affecting the arms, buttocks, and thighs, and thoracoabdominal adiposity also in a subcutaneous distribution. The other patient has minimal subcutaneous adiposity, perhaps even appearing to have lean, muscular arms and legs, and has negligible thoracoabdominal subcutaneous adiposity. However, inspection of the patient reveals a protuberant abdomen, implying substantial gain of adipose tissue in the visceral (intra-abdominal) compartment. Based on current understanding, the second of these 2 patients is more likely to have 1 or more manifestations of "toxic fat," known most commonly as the metabolic syndrome.[1] Notwithstanding, metabolic syndrome does occur in persons with BMIs considered "healthy" in the United States (23–25 kg/m^2). This emphasizes a truth that should not be lost on health practitioners in the United States–that an individual patient can be "fat on the inside" while having a normal BMI. What is important to understand and to communicate to patients is that expansion of visceral adiposity is associated with profound changes in signal hormones (adipokines) secreted by these adipocytes. This is at least associated with disease states, and, arguably, might be causative.[2]

THE METABOLIC SYNDROME CONCEPT

In 1988, Reaven coined the term "syndrome x" to conceptually link hypertension, T2DM, and increased propensity for atherosclerotic events to the putative underlying cause of insulin resistance.[3] The cardiology community already had a "syndrome x" for the phenomenon of angina pectoris associated with angiographically normal coronary arteries,[4] so Reaven's original name evolved to "metabolic syndrome," which has synonyms "insulin resistance syndrome" and "cardiometabolic syndrome." Based on data from the 1999–2000 National Health and Nutrition Examination Survey, the prevalence of metabolic syndrome in the United States now exceeds 25% of the adult population.[5]

> *Based on data from the 1999–2000 National Health and Nutrition Examination Survey, the prevalence of metabolic syndrome in the United States now exceeds 25% of the adult population.*

Metabolic syndrome has several different definitions with regard to its components, including those of the US National Cholesterol Education Program-Adult Treatment Panel III (NCEP-ATP III),[6] the World Health Organization,[7] the American College of Endocrinology (which names the syndrome "insulin resistance syndrome"),[8] and the International Diabetes Foundation.[9] The most widely used criteria in the United States are those of NCEP-ATP III, which are: (1) waist circumference >102 cm (>40 inches) for men and >88 cm (>35 inches) for women, (2) serum triglycerides (TG) ≥150 mg/dL, (3) high-density lipoprotein cholesterol (HDL) <40 mg/dL for men and <50 mg/dL for women, (4) blood pressure (BP) ≥130/85 mm Hg, and (5) fasting glucose ≥100 mg/dL.[6] The viability of the metabolic syndrome as a syndrome is undercut by lack of universal agreement on how insulin resistance is measured or defined, and whether a one-root-cause approach to treatment is viable. In 2006, a joint position statement by the American Diabetes Association and the European Association for the Study of Diabetes discouraged recognition of metabolic syndrome as a

> *In the United States, the most widely used criteria for the metabolic syndrome are those of the National Cholesterol Education Program-Adult Treatment Panel III, which are: (1) waist circumference >102 cm (>40 inches) for men and >88 cm (>35 inches) for women, (2) serum triglycerides ≥150 mg/dL, (3) high-density lipoprotein cholesterol (HDL) <40 mg/dL for men and <50 mg/dL for women, (4) blood pressure ≥130/85 mm Hg, and (5) fasting glucose ≥100 mg/dL.*

distinct entity, based on the following considerations: (1) criteria for the diagnosis of metabolic syndrome are "ambiguous or incomplete" and "the rationale for thresholds are ill-defined," (2) the value of including T2DM in the definition is unclear, (3) the role of insulin resistance as the cause of the syndrome is "uncertain," (4) there is no clear basis for including or excluding other risk factors for CVD from the syndrome definition, (5) CVD risk attributable to the metabolic syndrome may be no greater than that attributable to individual components of the syndrome, and (6) treatment for metabolic syndrome is no different than treatment of component features.[10]

Physicians who feel overwhelmed and confused by such academic controversy–and their patients–should feel comforted by the fact that although the definition of metabolic syndrome may be controversial, the concept is entirely viable, at least as clinical shorthand for risk of CVD and other component diseases.[11] As evidence, it is worthwhile to consider evidence afforded by several studies. Abbasi identified a strong correlation between 9 "traditional" coronary heart disease (CHD) risk factors including age, systolic BP, diastolic BP, total cholesterol, TG, HDL, glucose tolerance (as assessed by a standard 75-g oral glucose tolerance test [OGTT]), BMI, and insulin resistance measured by sophisticated in vivo techniques.[12] Wilson followed a cohort of 3323 men and women for incident CVD (CHD, stroke, new claudication, and congestive heart failure) and CHD over a period of 8 years, and found that risk of both CVD (relative risk [RR]: 2.88 for men, 2.25 for women) and CHD (RR: 2.54 for men, 1.54 for women) was increased among those satisfying NCEP-ATP III criteria for metabolic syndrome.[13] Sundstrom followed a cohort of 2322 men in a 32-year-longitudinal study, and found that those meeting either NCEP-ATP III or World Health Organization criteria for metabolic syndrome had increased cardiovascular mortality (RR: 1.59).[14] Identification of metabolic syndrome has been shown to perform similar to the Framingham Risk Score calculation in estimating cardiovascular risk, including in cohort studies by Wannamethee (5128 patients),[15] Najarian (2097 patients),[16] and Lorenzo (2559 patients).[17]

"LEAN HEALTHY" VERSUS "FAT TOXIC" ADIPOCYTES: METABOLIC YIN AND YANG

Once thought to be simply a storage bin for fat, the adipocyte is now known to be an active endocrine organ producing hormones (adipokines) that modulate insulin sensitivity, lipid composition, systemic inflammatory state, and endothelial function (Figure 1).

> *Once thought to be simply a storage bin for fat, the adipocyte is now known to be an active endocrine organ producing hormones (adipokines) that modulate insulin sensitivity, lipid composition, systemic inflammatory state, and endothelial function.*

Healthy adipocytes produce an adipokine milieu promoting normal glucose and lipid utilization, thrombolysis, healthy endothelial function with normotension, reduced oxidative stress, and minimal inflammation. By contrast, "fat toxic" adipocytes produce an adipokine milieu promoting insulin resistance with impaired glucose tolerance (IGT) or frank T2DM, dyslipidemia, a prothrombotic state, and disordered endothelial function characterized by hypertension, increased oxidative stress, and increased inflammation. Understanding of the exact mechanisms whereby adipocytes convert from a predominantly "lean healthy" state to a "fat toxic" state is evolving, but current concepts include relative

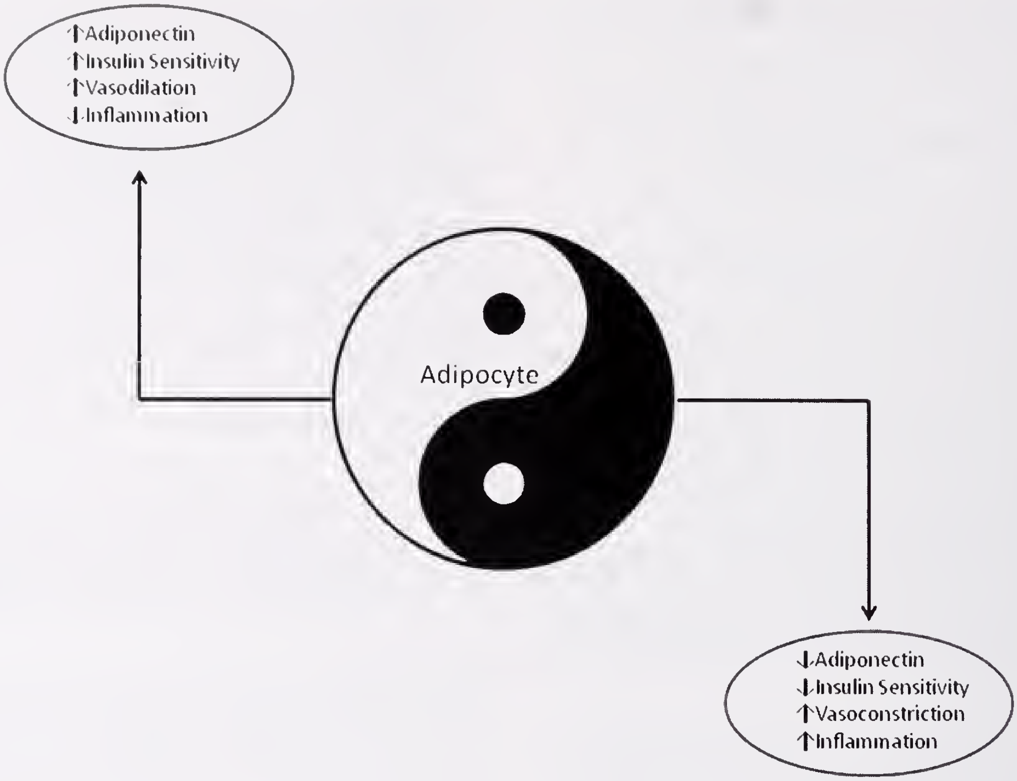


Fig 1. Metabolic "yin-yang" of the visceral adipocytes. In optimal health, "yin" predominates, characterized by an adiponectin-rich metabolic environment wherein insulin sensitivity is heightened, vascular inflammation is minimal, and arteriolar tone is relaxed, with a trend toward normotension. In metabolic syndrome, "yang" predominates: adiponectin secretion is low; alternate adipokines, including TNF-α and IL-6, promote a state of heightened vascular inflammation; and arteriolar tone is increased, with a trend toward hypertension. **Abbreviations:** IL-6, interleukin-6; TNF-α, tumor necrosis factor-alpha.

adipocyte hypoxia (related to cellular hypertrophy and hyperplasia), oxidative stress (from excessive TG accumulation and oxidation), and especially endoplasmic reticular stress, resulting in misfolding of cellular peptides (perhaps including adiponectin) and accumulation of unfolded peptide productions, which may generate a local inflammatory response to include increased generation of interleukin-6 (IL-6) and tumor necrosis factor-alpha (TNF-α).[18,19]

There is evidence that the hypothalamic-pituitary-adrenal axis is activated in metabolic syndrome, with a result that serum cortisol is higher than in unaffected patients, although frank Cushing syndrome is not present.[20] Overexpression of the enzyme 11β-hydroxysteroid dehydrogenase type 1 (11-β-HSD-1), which converts inert cortisone into bioactive cortisol, may explain the expansion of the visceral adipocyte compartment in disease; 11-β-HSD-1 is present in visceral adipose, but not in subcutaneous adipose tissue. Excessive activity of 11-β-HSD-1 may generate hypercortisolism within the visceral adipose tissue, resulting in impaired insulin action locally. Perhaps this is one of the signals changing the adipokine milieu from "lean healthy" to "fat toxic." Heightened 11-β-HSD-1 activity in humans has been associated with central obesity, insulin resistance, hypertension, and dyslipidemia.[20,21] In addition to excessive local and possibly systemic cortisol, the sympathetic nervous system (SNS) is activated to a greater-than-normal degree, with resultant higher heart rate and BP,[22] and the renin-angiotensin-aldosterone system (RAAS) is also activated to a hypernormal extent.[23]

The "lean healthy" adipokine most prevalent in human serum is adiponectin.

The "lean healthy" adipokine most prevalent in human serum is adiponectin.

Adiponectin promotes normal glucose tolerance by inhibiting gluconeogenesis in the liver,[21] by stimulating glucose transporter type 4–mediated glucose uptake by muscle tissue,[2] and by reducing circulating free fatty acids (FFA).[24] Adiponectin exerts systemic anti-inflammatory effects by inhibiting synthesis and release of TNF-α and IL-6 from adipose tissue. Adiponectin also influences the systemic state of thrombosis and thrombolysis by inhibiting synthesis and release of plasminogen activator inhibitor-1 (PAI-1) from visceral adipocytes. In vascular endothelium, adiponectin stimulates production of nitric oxide (NO), which is a marker of "healthy" endothelial function,[25] and inhibits TNF-α–mediated endothelial synthesis of intercellular adhesion molecule-1 (ICAM-1) and vascular cell adhesion molecule-1.[26] Thus, in a state of optimal systemic metabolic health, the "lean healthy" adipocyte generates an adiponectin-dominant microenvironment wherein insulin action is optimal and consequently glucose and TG metabolism is normal, systemic inflammation is restrained, and an antithrombotic state is present, minimizing risk of coronary and arterial thrombosis. This state may be characterized as metabolic "yin."

In a state of optimal systemic metabolic health, the "lean healthy" adipocyte generates an adiponectin-dominant microenvironment wherein insulin action is optimal and consequently glucose and TG metabolism is normal, systemic inflammation is restrained, and an antithrombotic state is present, minimizing risk of coronary and arterial thrombosis. This state may be characterized as metabolic "yin."

The "fat toxic" adipocyte, by contrast, conceptually represents the metabolic "yang" state, which is characterized by reduced adiponectin expression and increased expression of TNF-α, IL-6, and PAI-1. TNF-α and IL-6, in fact, are known to directly inhibit adiponectin production.[27] Visceral adipose tissue generates 30% of whole-body circulating IL-6,[28] which is known to be an important cause of impaired insulin action in muscle and adipose tissue.[24] TNF-α is synergistic with IL-6 in inducing insulin resistance, probably by contributing to a general state of systemic inflammation.[24,29] Compounding these effects, in adiponectin-deficient, visceral adipocyte insulin-resistant tissues, insulin fails to optimally suppress release of FFA from the visceral compartment; downstream effects of increased FFA include increased hepatic gluconeogenesis, increased hepatic TG production as very-low-density lipoprotein, and impaired insulin-stimulated NO-mediated vasodilation in vascular tissues.[30] Adipocyte-derived TNF-α and IL-6 are secreted directly into the portal circulation, enabling their synergistic effects to increase hepatic production of fibrinogen and C-reactive protein (CRP), thereby augmenting the systemic prothrombotic and inflammatory state.[2] As the adipose compartment physically expands, the "lean healthy" adiponectin-dominant environment (metabolic "yin") is converted to the "fat toxic" adiponectin-poor, inflammatory cytokine–enriched microenvironment in which adipocyte-derived TNF-α, IL-6, and PAI-1 replace adiponectin as the predominant hormonal products (metabolic "yang"). Thus, a state of systemic insulin resistance, inflammation, and prothrombotic tendency is established.

As the adipose compartment physically expands, the "lean healthy" adiponectin-dominant environment (metabolic "yin") is converted to the "fat toxic" adiponectin-poor, inflammatory cytokine–enriched microenvironment in which adipocyte-derived TNF-α, IL-6, and PAI-1 replace adiponectin as the predominant hormonal products (metabolic "yang"), thus establishing a state of systemic insulin resistance, inflammation, and prothrombotic tendency.

Downstream effects of adiponectin, TNF-α, and IL-6 are expressed in vascular endothelium through the intracellular signal pathway nuclear factor-kappa beta (NF-$\kappa\beta$). Translocation of NF-$\kappa\beta$ from cytosol to nucleus amplifies gene transcription of TNF-α, IL-6, matrix metalloproteinases, and cellular adhesion molecules by endothelial cells, and results in increased generation of reactive oxygen species. TNF-α and IL-6 promote this translocation, whereas adiponectin opposes it. NF-$\kappa\beta$'s translocation to the cell nucleus can only occur when it dissociates from

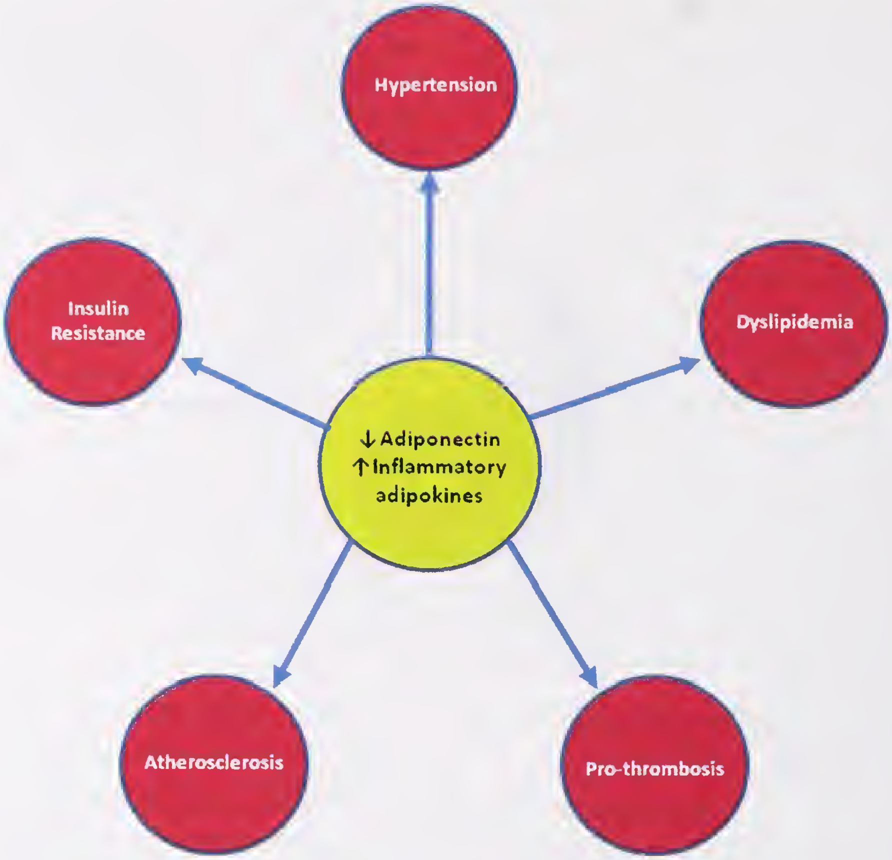


Fig 2. Revised concept of the "wheel and spokes" of metabolic syndrome. Insulin resistance has often been depicted as the hub, with multiple manifestations, including T2DM, hypertension, atherosclerosis, and a heightened systemic sate of inflammation and thrombosis. It may be that derangement of adipokinesis is actually the central event responsible for insulin resistance, and for other components of the metabolic syndrome. **Abbreviations:** T2DM, type 2 diabetes mellitus.

an inhibitory peptide, I$\kappa\beta$, to which it is bound in the cytosol. Dissociation is promoted by TNF-α, IL-6, and nonesterified free fatty acids, and inhibited by insulin and adiponectin.[28,31] In vascular endothelium, adiponectin stimulates NO-mediated vasodilation, impedes endothelin secretion and action to promote vasoconstriction, and inhibits macrophage conversion to atherogenic foam cells.[32]

Most of the adverse health consequences linked to obesity relate directly to the concept of metabolic "yang" and the "fat toxic" adipose compartment (primarily but not exclusively visceral adipose tissue) characterized by low adiponectin and high circulating TNF-α, IL-6, PAI-1, fibrinogen, and CRP; these include obesity,[33] essential hypertension,[34,35] IGT/prediabetes,[36–38] T2DM,[24,33,39,40] dyslipidemia (low HDL, hypertriglyceridemia, and small dense low-density lipoprotein cholesterol [LDL] particles),[41,42] acute coronary syndromes,[43–46] and chronic CHD (Figure 2).[37,47] TNF-α, IL-6, PAI-1, and high-sensitivity CRP (hs-CRP) serum levels have been strongly predictive of cardiovascular[48] and T2DM risk.[49] Clinical interventions that improve adiponectin activity also lower TNF-α, IL-6, PAI-1, fibrinogen, and CRP; regular aerobic exercise and weight reduction are the nonpharmacologic therapeutic lifestyle changes known to raise adiponectin production and improve risk profiles for CVD and T2DM.[50,51]

TREATMENTS TO PREVENT TYPE 2 DIABETES MELLITUS

As noted, one of the principal critiques offered by the joint American Diabetes Association/European Association for the Study of Diabetes position paper is a putative lack of evidence that treating metabolic syndrome independently of T2DM offers long-term benefits discrete from treatment of DM itself–in other words, appropriate care of the components of metabolic syndrome may produce outcomes

Table 1. *Type 2 Diabetes Mellitus Prevention Trials.*

Trial	No. of Subjects	Intervention	Duration, yr	Risk Reduction Versus Control Group, %
Diabetes Prevention Program[52]	3234	Diet + exercise	3	58
Finnish Diabetes Prevention Study[53]	522	Diet + exercise	3	58
Da Qing IGT and Diabetes Study[54]	577	Diet + exercise	6	43
Diabetes Prevention Program[52]	3234	Metformin	3	31
TRIPOD[77]	133	Troglitazone	2.5	55
DREAM[56]	2365	Rosiglitazone	3	60
STOP-NIDDM[78]	714	Acarbose	3	25

Abbreviations: DREAM, Diabetes Reduction Assessment with Ramipril and Rosiglitazone Medication; IGT, impaired glucose tolerance; STOP-NIDDM, Study to Prevent Non-Insulin-Dependent Diabetes Mellitus; TRIPOD, Troglitazone in the Prevention of Diabetes.

equivalent to syndrome-based treatment. A counterview is that there actually are several clinical trials that examine the efficacy of simple, root-cause-oriented lifestyle modifications or pharmacotherapy on both incident T2DM and cardiovascular endpoints. Examination of several of these trials (Table 1) is pertinent for the clinician interested in advising patients perceived to be at risk.

The DPP was a US multicenter, randomized clinical trial that enrolled 3234 patients to one of 3 treatment arms. Eligibility was determined by positive family history of T2DM and by identification of IGT on a standard 75-g OGTT. One arm of the study was a placebo group; one was a group that entered into an intensely supervised lifestyle-modification program to achieve ≥150 minutes per week of aerobic physical activity and achieve and maintain a 7% weight loss; and the third arm was a group prescribed metformin, 850 mg twice daily, with no prescribed lifestyle modification. Incidence of T2DM was reduced by >50% in the lifestyle-intervention group and by about one-third in the metformin treatment group, versus the placebo group. Numbers needed to treat to prevent 1 case of T2DM were 6.9 for lifestyle intervention and 13.9 for metformin.[52] Finnish[53] and Chinese[54] studies of similar design produced comparable results to the US study with regard to the impact of lifestyle modification on incident T2DM. The commitment to regular exercise and weight reduction is very difficult to sustain in clinical practice, and unfortunately, none of the studies to date have assessed whether or not medication therapy plus lifestyle interventions produce synergistic benefit, although it is logical to believe that it might. With the availability of low-cost, generic metformin, it seems very reasonable to prescribe metformin to patients at high risk for T2DM, based on the findings of the DPP.

In the Diabetes Prevention Program, a US multicenter, randomized clinical trial involving 3234 patients, the incidence of type 2 diabetes mellitus was reduced by >50% in the lifestyle-intervention group and by about one-third in the metformin treatment group, versus the placebo group. Numbers needed to treat to prevent 1 case of type 2 diabetes mellitus were 6.9 for lifestyle intervention and 13.9 for metformin.[52]

The Diabetes Reduction Assessment with Ramipril and Rosiglitazone Medication (DREAM) trial examined the efficacy of use of the thiazolidinedione (TZD) rosiglitazone in preventing onset of T2DM. The study included >5200 patients (60% female) from 21 countries who had IGT on a 75-g OGTT. Randomization was to either rosiglitazone 8 mg daily or to placebo; no lifestyle changes were enforced, although they were taught and encouraged. After 3 years of study, rosiglitazone-treated patients experienced an absolute risk reduction of 14.4% for the development of T2DM.[55] More recent findings implicating rosiglitazone with excess cardiovascular mortality[56] undercut enthusiasm for use of this specific agent in the prevention of T2DM, but it is possible that newer-generation TZDs will be free of this association.

LIFESTYLE MODIFICATIONS IN THE TREATMENT OF HYPERTENSION

BP can be lowered by 4–6 mm Hg systolic and 2–4 mm Hg diastolic in association with regular aerobic exercise, as evidenced by 2 meta-analyses of studies including large numbers of patients.[57,58] The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC 7) estimated a reduction in systolic BP of 5–20 mm Hg for every 10-kg weight loss, and regular aerobic exercise (30 minutes walking at a brisk pace, or equivalent) most days of the week was also estimated to produce a reduction of 4–9 mm Hg in systolic BP.[59] If not already hypertensive, many persons with obesity will meet the JNC 7 criteria for prehypertension, with BP 120–139 mm Hg systolic over 80–89 mm Hg diastolic. Lifestyle modifications, including diet and regular aerobic exercise with a goal of attaining and maintaining normal BMI, are primary therapies advocated by JNC 7 for management of prehypertension.[59]

> *As evidenced by 2 meta-analyses of studies including large numbers of patients, blood pressure can be lowered by 4–6 mm Hg systolic and 2–4 mm Hg diastolic in association with regular aerobic exercise.*

TREATMENT OF DYSLIPIDEMIA

Many obese patients judged to have metabolic syndrome will have 3 or more risk factors for CVD, and accordingly will fall into a risk/benefit category favoring aggressive treatment of dyslipidemia. The Anglo-Scandinavian Cardiac Outcomes Trial-Lipid Lowering Arm (ASCOT-LLA) study examined the impact of statin therapy in non-T2DM patients aged 40–79 years with hypertension and 3 other cardiovascular risk factors. More than 10,300 patients were randomized to either atorvastatin 10 mg daily or placebo; primary endpoints of the study were nonfatal myocardial infarction and fatal CHD events. At baseline, subjects had "average" lipid profiles: mean serum LDL was 132 mg/dL, HDL was 50 mg/dL, and TG were 150 mg/dL. The study was halted after 3 years due to an absolute risk reduction of 1.1% for primary endpoints in the group receiving atorvastatin (number needed to treat to prevent 1 endpoint event: 90.9).[60] The Helsinki Heart Study was a primary CHD prevention trial that randomized >4000 men aged 40–55 years in Finland to receive either placebo or the fibrate drug gemfibrozil 600 mg twice daily. At the time of randomization, mean serum lipids were LDL 189 mg/dL, HDL 48 mg/dL, and TG 179 mg/dL. Primary endpoints were fatal and nonfatal myocardial infarction and total cardiovascular death. In men with HDL <42 mg/dL and TG >200 mg/dL, there was a 40% reduction in primary endpoints, as compared with placebo-treated patients.[61] With efficacy established for both statins and fibrates in the prevention of CHD, physicians and patients may wish to consider one or the other on the basis of the major lipid abnormality: for high LDL and normal TG, a statin probably is more rational, and for low LDL, low HDL, high TG, and high non-HDL cholesterol (calculated as total cholesterol − HDL), initial use of a fibrate drug seems reasonable.

> *With efficacy established for both statins and fibrates in the prevention of coronary heart disease, physicians and patients may wish to consider one or the other on the basis of the major lipid abnormality: for high LDL and normal TG, a statin probably is more rational, and for low LDL, low HDL, high TG, and high non-HDL cholesterol (calculated as total cholesterol − HDL), initial use of a fibrate drug seems reasonable.*

TREATMENT OF HYPERTENSION AND PREHYPERTENSION

A brief summary of what is known about the pathophysiology of hypertension and prehypertension in patients with obesity is pertinent and should be of interest to patients and their physicians. The low serum adiponectin and high levels of inflammatory adipokines are associated with impaired generation of endothelium-derived NO, an important vasodilatory substance. Additionally, production of endothelium-derived endothelin-1, a potent vasoconstrictor, is increased. The SNS is activated by the adiponectin-poor, inflammatory adipokine–enriched milieu, resulting in high resting heart rate and increased norepinephrine-mediated vasoconstriction. The RAAS is likewise activated, possibly because of renal afferent arteriolar constriction, partially because of increased sympathetic signaling to the juxtaglomerular apparatus and in part by activation of mineralocorticoid receptors in renal tubular epithelium (and in other tissues) by cortisol, rather than aldosterone.[2] Finally, another unique adipokine, leptin, is elevated in obesity, generally correlating with adipose mass. Leptin is a satiety-promoting hormone in the central nervous system (CNS), and for reasons not yet clear most human obesity is associated with CNS resistance to leptin signaling. In the kidney, leptin seems to stimulate local collagen synthesis and glomerular cellular expansion, and has been implicated in the focal glomerular sclerosis of obesity.[62] Its exact contribution to hypertension in obese individuals is not entirely known, and is worthy of ongoing study. In summary, obesity is associated with subnormal endothelial-derived NO-mediated vasodilation, potentiation of endothelial-derived endothelin-1–mediated vasoconstriction, activation of the SNS, activation of the RAAS, and direct glomerular damage. Not surprisingly, hypertension is common in obesity and poses an opportunity to tailor lifestyle and medication therapy toward correction of root causes.

ANTIHYPERTENSIVE CHOICES IN HYPERTENSION AND PREHYPERTENSION

Because of their salutary effects on the RAAS and association of use with reduced risk of CVD, progressive renal disease, and even retinopathy, and because generic availability now makes the therapy low-cost,[7] angiotensin-converting enzyme inhibitors (ACEIs) are a logical first choice for treatment of hypertension in obese patients with metabolic syndrome. Among the many choices of antihypertensive agents for patients with metabolic syndrome, new information suggests a preferential role for angiotensin receptor blockers (ARBs). The ARB candasartan has been associated with reduced circulating ICAM-1 and improved endothelial function, as represented by reactive-hyperemia-induced forearm blood flow in patients with hypercholesterolemia.[63] Irbesartan and candasartan (but not losartan) have been shown to reduce plasma PAI-1.[64]

Because angiotensin-converting enzyme inhibitors have salutary effects on the renin-angiotensin-aldosterone system; are associated with reduced risk of cardiovascular disease, progressive renal disease, and even retinopathy; and are now available generically at low cost, they are a logical first choice for treatment of hypertension in obese patients with metabolic syndrome. New information also suggests a preferential role in this regard for angiotensin receptor blockers.

Previous concerns that thiazide diuretics (TTD) might actually cause T2DM have been largely assuaged by recent work using lower doses,[65–67] but the combination of diuretics and β-blockers is associated with increased risk for development of T2DM. Because the ACEI and ARB classes of drugs are especially tailored to pathophysiology of hypertension in obesity (and T2DM), a TTD should probably be reserved as second-line therapy, or for patients exhibiting fluid retention. There is now considerable interest in using the aldosterone antagonist spironolactone preferentially when a diuretic is needed; the disadvantage is the possibility of hyperkalemia when an ACEI or ARB is used in combination with spironolactone, and the greater financial cost of spironolactone versus TTDs.

The TZD pioglitazone is associated with an increase in LDL particle size from the small, dense (proatherogenic) toward the large, "fluffy" (less atherogenic) size.[68] There is evidence that TZDs may also improve BP in hypertensive patients. Rosiglitazone was associated with a reduction in BP from 138 ± 2 mm Hg systolic over 85 ± 2 mm Hg diastolic to 134 ± 2 mm Hg systolic over 80 ± 2 mm Hg diastolic[69] in patients already treated for hypertension, and a similar effect has been demonstrated for pioglitazone.[70] TZDs increase plasma adiponectin better than any other available drug therapy.[70–72] However, because of concerns about fluid retention, congestive heart failure, and possibly increased CVD risk, more needs to be known about TZDs before they can be recommended as a first-choice, early-use option in patients who obviously have increased CVD risk as a central feature of their syndrome.

SUMMARY OF RECOMMENDATIONS FOR TREATMENT OF PATIENTS WITH FEATURES OF METABOLIC SYNDROME

Diet (daily caloric restriction to <2000 kcal/day) and exercise therapy (≥150 minutes of brisk walking, or equivalent exercise, per week) can prevent or delay onset of T2DM and may reduce cardiovascular risk independent of T2DM risk, lower BP, and promote weight loss. Thus, these twin components of "therapeutic lifestyle changes" can be recommended enthusiastically to patients. An important lesson from the DPP is that adherence to lifestyle modification is enhanced with a support group and by direct professional supervision. This may be difficult for doctors and patients alike, but may be a component of health-system reform that all can strive to work for as a completely nonpartisan, health-promoting activity.

Metformin has been shown in a large randomized controlled trial (the DPP) to be associated with reduced incidence of T2DM, although inferior to diet and exercise in that regard. The salutary effects of TZDs on glycemia (reduced risk of T2DM in the DREAM trial), BP, and lipid composition make an agent of this pharmacologic class a tempting option for initial monotherapy in the patient with prediabetes (fasting glucose 100–125 mg/dL), prehypertension (for example, BP 135/84 mm Hg), and mild dyslipidemia (for example, male patient with HDL of 39 mg/dL and TG of 164 mg/dL). Buffering that enthusiasm is the high economic cost of TZDs, their association with long-term weight gain, and an unclear impact on CVD events.

As patients with metabolic syndrome are at increased cardiovascular risk, aggressive use of statins and/or fibrate drugs seems entirely reasonable. For patients with serum TG <250 mg/dL and LDL >100–130 mg/dL, a statin is the choice best justified by evidence from numerous clinical trials. If on-treatment serum TG remain significantly >250 mg/dL, and if HDL is <35 mg/dL, combining fenofibrate with a statin may also be reasonable. Although the routine use of combination statin and fenofibrate was not superior to statin alone in the Action to Control Risk Factors in Diabetes (ACCORD) trial with >10,000 patients with T2DM, there was possible benefit to the combination in the subgroup of patients with both elevated TG and low HDL.[73] Use of either statins or fibrates may be associated with health benefits not entirely attributable to their lipid-lowering effects. Combining statin drugs with diet has been shown to reduce CRP, ICAM-1, and fibrinogen,[74] and may be associated with a decreased risk of new-onset T2DM.[75] Fenofibrate is associated with improved insulin response and lower hs-CRP.[76]

CONCLUSION

Patients with metabolic syndrome are at risk for developing multiple disease processes, including T2DM, hypertension, dyslipidemia, and CVD, which may be concurrently in play. Therapies targeting the root cause of the problem—putatively, a switch from metabolic "yin" to metabolic "yang" in the adipose compartment—may benefit all concomitant diseases, save healthcare costs, and, not least of all, make the patient's self-esteem and ownership of health all the greater.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Bays HE. "Sick fat," metabolic disease, and atherosclerosis. *Am J Med* 2009; 122(suppl 1): S26–S37.
2. Gable DR, Hurel SJ, Humphries SE. Adiponectin and its gene variants as risk factors for insulin resistance, the metabolic syndrome and cardiovascular disease. *Atherosclerosis* 2006; 188: 231–244.
3. Reaven GM. Banting lecture 1988: role of insulin resistance in human disease. *Diabetes* 1988; 37: 1595–1607.
4. Kaski JC, Elliott PM. Angina pectoris and normal coronary arteriograms: clinical presentation and hemodynamic characteristics. *Am J Cardiol* 1995; 76: 35D–42D.
5. Ford ES, Giles WH, Mokdad AH. Increasing prevalence of the metabolic syndrome among U.S. adults. *Diabetes Care* 2004; 27: 2444–2449.
6. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Final Report. *Circulation* 2002; 106: 3143.
7. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation [see comment]. *Diabet Med* 1998; 15: 539–553.
8. Einhorn D, Reaven GM, Cobin RH, et al. American College of Endocrinology position statement on the insulin resistance syndrome. *Endocr Pract* 2003; 9: 237–252.
9. Alberti KG, Zimmet P, Shaw J. Metabolic syndrome—a new world-wide definition. A Consensus Statement from the International Diabetes Federation. *Diabet Med* 2006; 23: 469–480.
10. Kahn R, Buse J, Ferrannini E, Stern M. The metabolic syndrome: time for a critical appraisal: joint statement from the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care* 2005; 28: 2289–2304.
11. Haffner SM. Abdominal adiposity and cardiometabolic risk: do we have all the answers? *Am J Med* 2007; 120(9 suppl 1): S10–S17.
12. Abbasi F, Brown BW, Lamendola C, et al. Relationship between obesity, insulin resistance, and coronary heart disease risk. *J Am Coll Cardiol* 2002; 40: 937–943.
13. Wilson PW, D'Agostino RB, Parise H, et al. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation* 2005; 112: 3066–3072.

14. Sundstrom J, Riserus U, Byberg L, et al. Clinical value of the metabolic syndrome for long term prediction of total and cardiovascular mortality: prospective, population based cohort study. *BMJ* 2006; 332: 878–882.
15. Wannamethee SGP, Shaper AGF, Lennon LM, Morris RWP. Metabolic syndrome vs Framingham risk score for prediction of coronary heart disease, stroke, and type 2 diabetes mellitus. *Arch Intern Med* 2005; 165: 2644–2650.
16. Najarian RM, Sullivan LM, Kannel WB, et al. Metabolic syndrome compared with type 2 diabetes mellitus as a risk factor for stroke: the Framingham Offspring Study. *Arch Intern Med* 2006; 166: 106–111.
17. Lorenzo C, Williams K, Hunt KJ, Haffner SM. The National Cholesterol Education Program-Adult Treatment Panel III, International Diabetes Federation, and World Health Organization definitions of the metabolic syndrome as predictors of incident cardiovascular disease and diabetes. *Diabetes Care* 2007; 30: 8–13.
18. Ozcan U, Cao Q, Yilmaz E, et al. Endoplasmic reticulum stress links obesity, insulin action, and type 2 diabetes. *Science* 2004; 306: 457–461.
19. Gregor MF, Hotamisligil GS. Thematic review series: Adipocyte Biology. Adipocyte stress: the endoplasmic reticulum and metabolic disease. *J Lipid Res* 2007; 48: 1905–1914.
20. Anagnostis P, Athyros VG, Tziomalos K, et al. Clinical review: the pathogenetic role of cortisol in the metabolic syndrome: a hypothesis. *J Clin Endocrinol Metab* 2009; 94: 2692–2701.
21. Ahima RS. Adipose tissue as an endocrine organ. *Obesity* 2006; 14: 242–249.
22. Grassi G, Arenare F, Quarti-Trevano F, et al. Heart rate, sympathetic cardiovascular influences, and the metabolic syndrome. *Prog Cardiovasc Dis* 2009; 52: 31–37.
23. Sowers JR, Whaley-Connell A, Epstein M. Narrative review: the emerging clinical implications of the role of aldosterone in the metabolic syndrome and resistant hypertension. *Ann Intern Med* 2009; 150: 776–783.
24. Pittas AG, Joseph NA, Greenberg AS. Adipocytokines and insulin resistance. *J Clin Endocrinol Metab* 2004; 89: 447–452.
25. Chen H, Montagnani M, Funahashi T, et al. Adiponectin stimulates production of nitric oxide in vascular endothelial cells. *J Biol Chem* 2003; 278: 45021–45026.
26. Ouchi N, Kihara S, Arita Y, et al. Novel modulator for endothelial adhesion molecules: adipocyte-derived plasma protein adiponectin. *Circulation* 1999; 100: 2473–2476.
27. Fasshauer M, Kralisch S, Klier M, et al. Adiponectin gene expression and secretion is inhibited by interleukin-6 in 3T3-L1 adipocytes. *Biochem Biophys Res Commun* 2003; 301: 1045–1050.
28. Tataranni PA, Ortega E. A burning question: does an adipokine-induced activation of the immune system mediate the effect of overnutrition on type 2 diabetes? *Diabetes* 2005; 54: 917–927.
29. Koh KK, Han SH, Quon MJ. Inflammatory markers and the metabolic syndrome: insights from therapeutic interventions. *J Am Coll Cardiol* 2005; 46: 1978–1985.
30. Jensen MD. Role of body fat distribution and the metabolic complications of obesity. *J Clin Endocrinol Metab* 2008; 93(11 suppl 1): S57–S63.
31. Ouchi N, Kihara S, Arita Y, et al. Adiponectin, an adipocyte-derived plasma protein, inhibits endothelial NF-κB signaling through a cAMP-dependent pathway. *Circulation* 2000; 102: 1296–1301.
32. Han SH, Quon MJ, Kim JA. Koh KK. Adiponectin and cardiovascular disease: response to therapeutic interventions. *J Am Coll Cardiol* 2007; 49: 531–538.
33. Weyer C, Funahashi T, Tanaka S, et al. Hypoadiponectinemia in obesity and type 2 diabetes: close association with insulin resistance and hyperinsulinemia. *J Clin Endocrinol Metab* 2001; 86: 1930–1935.
34. Iwashima Y, Katsuya T, Ishikawa K, et al. Hypoadiponectinemia is an independent risk factor for hypertension. *Hypertension* 2004; 43: 1318–1323.
35. Adamczak M, Wiecek A, Funahashi T, et al. Decreased plasma adiponectin concentration in patients with essential hypertension. *Am J Hypertens* 2003; 16: 72–75.
36. Choi KM, Lee J, Lee KW, et al. The associations between plasma adiponectin, ghrelin levels and cardiovascular risk factors. *Eur J Endocrinol* 2004; 150: 715–718.
37. Higashiura K, Ura N, Ohata J, et al. Correlations of adiponectin level with insulin resistance and atherosclerosis in Japanese male populations. *Clin Endocrinol* 2004; 61: 753–759.
38. Tschritter O, Fritsche A, Thamer C, et al. Plasma adiponectin concentrations predict insulin sensitivity of both glucose and lipid metabolism. *Diabetes* 2003; 52: 239–243.
39. Lindsay RS, Funahashi T, Hanson RL, et al. Adiponectin and development of type 2 diabetes in the Pima Indian population. *Lancet* 2002; 360: 57.
40. Daimon M, Oizumi T, Saitoh T, et al. Decreased serum levels of adiponectin are a risk factor for the progression to type 2 diabetes in the Japanese population: the Funagata study. *Diabetes Care* 2003; 26: 2015–2020.
41. Hulthe J, Hultén LM, Fagerberg B. Low adipocyte-derived plasma protein adiponectin concentrations are associated with the metabolic syndrome and small dense low-density lipoprotein particles: atherosclerosis and insulin resistance study. *Metabolism* 2003; 52: 1612–1614.
42. Matsubara M, Maruoka S, Katayose S. Decreased plasma adiponectin concentrations in women with dyslipidemia. *J Clin Endocrinol Metab* 2002; 87: 2764–2769.
43. Pischon T, Girman CJ, Hotamisligil GS, et al. Plasma adiponectin levels and risk of myocardial infarction in men. *JAMA* 2004; 291: 1730–1737.
44. Kumada M, Kihara S, Sumitsuji S, et al. Association of hypoadiponectinemia with coronary artery disease in men. *Arterioscler Thromb Vasc Biol* 2003; 23: 85–89.
45. Zoccali C, Mallamaci F, Tripepi G, et al. Adiponectin, metabolic risk factors, and cardiovascular events among patients with end-stage renal disease. *J Am Soc Nephrol* 2002; 13: 134–141.
46. Schulze MB, Shai I, Rimm EB, et al. Adiponectin and future coronary heart disease events among men with type 2 diabetes. *Diabetes* 2005; 54: 534–539.
47. Okamoto Y, Kihara S, Funahashi T, et al. Adiponectin: a key adipocytokine in metabolic syndrome. *Clin Sci (Lond)* 2006; 110: 267–278.
48. Duncan BB, Schmidt MI, Pankow JS, et al. Low-grade systemic inflammation and the development of type 2 diabetes: the Atherosclerosis Risk in Communities Study. *Diabetes* 2003; 52: 1799–1805.
49. Hu FB, Meigs JB, Li TY, et al. Inflammatory markers and risk of developing type 2 diabetes in women. *Diabetes* 2004; 53: 693–700.

50. Esposito KM, Pontillo AM, Di Palo C, et al. Effect of weight loss and lifestyle changes on vascular inflammatory markers in obese women: a randomized trial. *JAMA* 2003; 289: 1799–1804.
51. Yang WS, Lee WJ, Funahashi T, et al. Weight reduction increases plasma levels of an adipose-derived anti-inflammatory protein, adiponectin. *J Clin Endocrinol Metab* 2001; 86: 3815–3819.
52. Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002; 346: 393–403.
53. Tuomilehto J, Lindstrom J, Eriksson JG, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *N Engl J Med* 2001; 344: 1343–1350.
54. Pan XR, Li GW, Hu YH, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the Da Qing IGT and Diabetes Study. *Diabetes Care* 1997; 20: 537–544.
55. DREAM Trial Investigators. Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial. *Lancet* 2006; 368: 1096–1105.
56. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes [published correction appears in *N Engl J Med* 2007; 357: 100]. *N Engl J Med* 2007; 356: 2457–2471.
57. Whelton SP, Chin A, Xin X, He J. Effect of aerobic exercise on blood pressure: a meta-analysis of randomized, controlled trials. *Ann Intern Med* 2002; 136: 493–503.
58. Kelley GA, Kelley KS, Tran ZV. Walking and resting blood pressure in adults: a meta-analysis. *Prevent Med* 2001; 33(2 part 1): 120–127.
59. Chobanian AV, Bakris GL, Black HR, et al. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure [see comment]. *Hypertension* 2003; 42: 1206–1252.
60. Sever PS, Dahlof B, Poulter NR, et al. Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial–Lipid Lowering Arm (ASCOT-LLA): a multicentre randomised controlled trial [see comment]. *Lancet* 2003; 361: 1149–1158.
61. Manninen V, Tenkanen L, Koskinen P, et al. Joint effects of serum triglyceride and LDL cholesterol and HDL cholesterol concentrations on coronary heart disease risk in the Helsinki Heart Study. Implications for treatment [see comment]. *Circulation* 1992; 85: 37–45.
62. Reisin E, Jack AV. Obesity and hypertension: mechanisms, cardio-renal consequences, and therapeutic approaches. *Med Clin North Am* 2009; 93: 733–751.
63. Wassmann S, Hilgers S, Laufs U, et al. Angiotensin II type 1 receptor antagonism improves hypercholesterolemia-associated endothelial dysfunction. *Arterioscler Thromb Vasc Biol* 2002; 22: 1208–1212.
64. Koh KK, Han SH, Chung WJ, et al. Comparison of effects of losartan, irbesartan, and candesartan on flow-mediated brachial artery dilation and on inflammatory and thrombolytic markers in patients with systemic hypertension. *Am J Cardiol* 2004; 93: 1432–1435.
65. Gress TW, Nieto FJ, Shahar E, et al.; The Atherosclerosis Risk in Communities Study. Hypertension and antihypertensive therapy as risk factors for type 2 diabetes mellitus. *N Engl J Med* 2000; 342: 905–912.
66. Gurwitz JH, Bohn RL, Glynn RJ, et al. Antihypertensive drug therapy and the initiation of treatment for diabetes mellitus. *Ann Intern Med* 1993; 118: 273–278.
67. Savage PJ, Pressel SL, Curb JD, et al. Influence of long-term, low-dose, diuretic-based, antihypertensive therapy on glucose, lipid, uric acid, and potassium levels in older men and women with isolated systolic hypertension: the Systolic Hypertension in the Elderly Program. *Arch Intern Med* 1998; 158: 741–751.
68. Winkler K, Konrad T, Füllert S, et al. Pioglitazone reduces atherogenic dense LDL particles in nondiabetic patients with arterial hypertension: a double-blind, placebo-controlled study. *Diabetes Care* 2003; 26: 2588–2594.
69. Raji A, Seely EW, Bekins SA, et al. Rosiglitazone improves insulin sensitivity and lowers blood pressure in hypertensive patients. *Diabetes Care* 2003; 26: 172–178.
70. Hirose H, Kawai T, Yamamoto Y, et al. Effects of pioglitazone on metabolic parameters, body fat distribution, and serum adiponectin levels in Japanese male patients with type 2 diabetes. *Metabolism* 2002; 51: 314–317.
71. Maeda N, Takahashi M, Funahashi T, et al. PPARγ ligands increase expression and plasma concentrations of adiponectin, an adipose-derived protein. *Diabetes* 2001; 50: 2094–2099.
72. Yu JG, Javorschi S, Hevener AL, et al. The effect of thiazolidinediones on plasma adiponectin levels in normal, obese, and type 2 diabetic subjects. *Diabetes* 2002; 51: 2968–2974.
73. The ACCORD Study Group. Effects of combination lipid therapy in type 2 diabetes mellitus. *N Engl J Med* 2010; 362: 1563–1574.
74. Koh KK, Son JW, Ahn JY, et al. Vascular effects of diet and statin in hypercholesterolemic patients. *Int J Cardiol* 2004; 95: 185–191.
75. Freeman DJ, Norrie J, Sattar N, et al. Pravastatin and the development of diabetes mellitus: evidence for a protective treatment effect in the West of Scotland Coronary Prevention Study. *Circulation* 2001; 103: 357–362.
76. Koh KK, Han SH, Quon MJ, et al. Beneficial effects of fenofibrate to improve endothelial dysfunction and raise adiponectin levels in patients with primary hypertriglyceridemia. *Diabetes Care* 2005; 28: 1419–1424.
77. Buchanan TA, Xiang AH, Peters RK et al. Preservation of pancreatic beta-cell function and prevention of type 2 diabetes by pharmacological treatment of insulin resistance in high-risk Hispanic women. *Diabetes* 2002; 51: 2796–2803.
78. Chiasson JL, Josse RG, Gomis R, et al. Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomised trial. *Lancet* 2002; 359: 2072–2077.

Mount Sinai Journal of Medicine 77:511–523, 2010

Insulin Resistance in Obesity as the Underlying Cause for the Metabolic Syndrome

Emily J. Gallagher, MB, MRCPI,[1] Derek LeRoith, MD, PhD,[1] and Eddy Karnieli, MD[2]

[1]Mount Sinai School of Medicine, New York, NY
[2]Rambam Medical Center, Rappaport Faculty of Medicine, Technion, Haifa, Israel

OUTLINE

OBESITY AND THE METABOLIC SYNDROME
PREVALENCE
PATHOGENESIS
  Normal Insulin Action
  Normal Insulin Signaling
  Factors Involved in Insulin Resistance
METABOLIC EFFECTS OF INSULIN RESISTANCE
  Hypertension and Insulin Resistance
  Dyslipidemia and Insulin Resistance
  Hyperglycemia and Insulin Resistance
CONCLUSION

ABSTRACT

The metabolic syndrome affects more than a third of the US population, predisposing to the development of type 2 diabetes and cardiovascular disease. The 2009 consensus statement from the International Diabetes Federation, American Heart Association, World Heart Federation, International Atherosclerosis Society, International Association for the Study of Obesity, and the National Heart, Lung, and Blood Institute defines the metabolic syndrome as 3 of the following elements: abdominal obesity, elevated blood pressure, elevated triglycerides, low high-density lipoprotein cholesterol, and hyperglycemia. Many factors contribute to this syndrome, including decreased physical activity, genetic predisposition, chronic inflammation, free fatty acids, and mitochondrial dysfunction. Insulin resistance appears to be the common link between these elements, obesity and the metabolic syndrome. In normal circumstances, insulin stimulates glucose uptake into skeletal muscle, inhibits hepatic gluconeogenesis, and decreases adipose-tissue lipolysis and hepatic production of very-low-density lipoproteins. Insulin signaling in the brain decreases appetite and prevents glucose production by the liver through neuronal signals from the hypothalamus. Insulin resistance, in contrast, leads to the release of free fatty acids from adipose tissue, increased hepatic production of very-low-density lipoproteins and decreased high-density lipoproteins. Increased production of free fatty acids, inflammatory cytokines, and adipokines and mitochondrial dysfunction contribute to impaired insulin signaling, decreased skeletal muscle glucose uptake, increased hepatic gluconeogenesis, and β cell dysfunction, leading to hyperglycemia. In addition, insulin resistance leads to the development of hypertension by impairing vasodilation induced by nitric oxide. In this review, we discuss normal insulin signaling and the mechanisms by which insulin resistance contributes to the development of the metabolic syndrome. ***Mt Sinai J Med 77:511–523, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: diabetes mellitus, dyslipidemia, hypertension, insulin resistance, metabolic syndrome.

Address Correspondence to:

Eddy Karnieli
Rambam Medical Center
Haifa, Israel
Email: eddy@tx.technion.ac.il

The metabolic syndrome is a collection of related risk factors that predispose to the development of type 2 diabetes (T2DM) and cardiovascular disease (CVD). The term "metabolic syndrome" was coined by Haller in 1977 to describe the collection of risk factors (obesity, diabetes, hyperlipidemia, hyperuricemia, and steatic hepatosis) that were found in a German population to increase the risk of developing ischemic heart disease.[1] The complications of obesity have been recognized since ancient times; Hippocrates noted that "sudden death is more common in those that are naturally fat than lean," possibly describing cardiac disease

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20212

© 2010 Mount Sinai School of Medicine

associated with obesity.[2] The 17th-century Dutch surgeon Nicholas Tulp described the case of a patient with obesity, hypertriglyceridemia, and bleeding, linking obesity with lipid abnormalities and hypertension.[3] Subsequently, in the 18th century, the Italian anatomist Giovanni Battista Morgagni recorded his postmortem findings in 2 patients with visceral adiposity and reported evidence of hypertension, dyslipidemia, and hyperuricemia along with sleep apnea, heart failure, and cerebrovascular disease.[4] The original description of the metabolic syndrome is frequently attributed to the Swedish physician Kylin, who in 1923 described a syndrome of hypertension, dyslipidemia, and hyperuricemia in addition to hyperglycemia.[5] In his Banting lecture in 1988, Reaven described hyperinsulinemia and insulin resistance (IR) as the common thread linking these manifestations and referred to this syndrome as "Syndrome X."[6]

Table 1. *Criteria for Diagnosis of the Metabolic Syndrome.*

Increased WC	Cutoff*
Elevated TG or on treatment for elevated TG	≥150 mg/dL
Decreased HDL or on treatment for decreased HDL	Men <40 mg/dL; Women <50 mg/dL
Elevated BP or on treatment for hypertension	Systolic BP ≥130 mm Hg or; diastolic BP ≥85 mm Hg
Elevated FBG	≥100 mg/dL

Abbreviations: BP, blood pressure; FBG, fasting blood glucose; HDL, high-density lipoprotein cholesterol; IDF, International Diabetes Federation; NHLBI, National Heart, Lung, and Blood Institute; TG, triglycerides; WC, waist circumference.

* Based on population-specific or country-specific criteria; WC cutoff varies by ethnicity and country. IDF male cutoff measurements are ≥94 cm for Europoid, Middle Eastern, and sub-Saharan African populations; and ≥90 cm for Asian and Central and South American populations. IDF female cutoff measurements are ≥80 cm for all populations. NHLBI recommended cutoff is ≥102 cm for US males and ≥88 cm for US females.[12]

The first formal definition of the metabolic syndrome was proposed by the World Health Organization in 1998; however, this definition was thought to be impractical by many, as it included a hyperinsulinemic euglycemic clamp study as one of the diagnostic criteria for IR.[7] Various other societies including the European Group for the Study of Insulin Resistance, the National Cholesterol Education Program Adult Treatment Panel III (ATP III), the American Association of Clinical Endocrinologists, the International Diabetes Federation (IDF), the American Heart Association (AHA), and the National Heart, Lung, and Blood Institute (NHLBI), have published different definitions of the metabolic syndrome since that time.[8–11] The conflicting definitions have caused uncertainty among physicians as to how to identify patients with the metabolic syndrome and inconsistencies in epidemiological research, depending on what criteria are used. In 2009, the IDF, AHA, NHLBI, World Heart Federation, International Atherosclerosis Society, and International Association for the Study of Obesity published a joint statement with a harmonized definition of the metabolic syndrome.[12] According to this statement, any 3 of the following 5 risk factors constitutes a diagnosis of the metabolic syndrome (Table I): enlarged waist circumference (WC) with different population-specific and country-specific criteria; elevated triglycerides (TG) of ≥1.7 mmol/L (150 mg/dL); reduced high-density lipoprotein (HDL) cholesterol of <1.03 mmol/L (40 mg/dL) in men and <1.29 mmol/L (50 mg/dL) in women; elevated blood pressure, with a systolic blood pressure ≥130 mm Hg or a diastolic blood pressure ≥85 mm Hg; and elevated fasting glucose ≥5.6 mmol/L (100 mg/dL), with the inclusion of those individuals using medication to treat hypertriglyceridemia, decrease HDL, hypertension, or hyperglycemia.[12] The clinical significance of the metabolic syndrome has been examined in prospective studies. A meta-analysis of these studies reported that it carries an increased relative risk (RR) of developing T2DM of 2.99, along with a RR of 1.65 of developing CVD and a RR of 1.27 for all-cause mortality.[13]

> *A meta-analysis of prospective studies concerning the metabolic syndrome has shown that it carries an increased relative risk of developing type 2 diabetes mellitus of 2.99, along with a relative risk of 1.65 of developing cardiovascular disease and a relative risk of 1.27 for all-cause mortality.*

Significant controversy surroundes the metabolic syndrome, its definitions, and clinical value. Those who oppose its existence as a clinical syndrome question the evidence that has led to the cutoff values that define the metabolic syndrome. They object to it being called a syndrome, given that it does not have a single clear etiology. In addition, they argue that the syndrome adds little to the traditional Framingham risk factors for predicting the development of CVD and suggest that a fasting glucose may offer as much information about the risk of developing subsequent diabetes as the more complex components of the metabolic syndrome. Finally, they propose that it does not enhance the clinical management of these patients, in that the treatment guidelines for the metabolic syndrome begin with diet and exercise, so they query whether a physician would not offer the same advice to any patient with any one of the individual components of the metabolic syndrome.[14,15] However, some studies have suggested that up to 40% of obese individuals may be metabolically normal and may not be at greater risk of developing T2DM or CVD.[16,17] Therefore, diagnosing the metabolic syndrome and giving it a label emphasizes the importance of assessing multiple risk factors in patients and highlights the importance of lifestyle interventions.[18] One recent large study, the National Health and Nutrition Examination Survey III, reported that the metabolically normal obese person may be a rarer phenomenon than previously reported and appears to carry an increased risk of all-cause mortality similar to that of the metabolically abnormal obese individual.[19]

OBESITY AND THE METABOLIC SYNDROME

Obesity is identified by measuring WC, rather than body-mass index (BMI), when assessing a patient for the presence of the metabolic syndrome. WC correlates with visceral adiposity and IR and has a stronger association with the development of T2DM and CVD than does BMI.[12,20,21] The 2009 definition of the metabolic syndrome (Table 1) has been updated to include ethnic-specific WC cutoff values to diagnose obesity, as it is now recognized that certain ethnic groups, particularly South Asian populations, have greater amounts of visceral adiposity for given WC measurements compared with European populations.[10,12,22] There is still uncertainty regarding the appropriate cutoff values for diagnosing abdominal obesity in many ethnic groups, due to a lack of sufficient epidemiological data on WC and metabolic risk in many populations.[11,12]

Compared with European populations, certain other ethnic groups, particularly South Asian populations, have greater amounts of visceral adiposity for given measurements of waist circumference.

The World Health Organization identifies two cutoff measurements of WC in men and women, the lower ($\geq$94 cm in men and $\geq$80 cm in women) being associated with an increased risk of metabolic complications, and the higher ($\geq$102 cm in men and $\geq$88 cm in women), which correlates with a BMI of $\sim$30 kg/m^2 in men, associated with an even greater risk. The higher cutoff measurements are used by the AHA/NHLBI, Health Canada, and the European cardiovascular societies to identify abdominal obesity in US, Canadian, and European populations, respectively.[12] However, the IDF suggests the lower cutoff should be used in Europoid, Middle Eastern, Mediterranean, and sub-Saharan African populations to define obesity. South Asian populations, Chinese, and Japanese populations, along with Central and South American populations, should be considered obese with a WC of $\geq$90 cm in men and $\geq$80 cm in women, according to the IDF.[10] The Chinese Cooperative Task Force, however, applies even lower values, of $\geq$85 cm in men and $\geq$80 cm in women.[12] The Japanese Obesity Society is the only society that applies a higher cutoff in women than men ($\geq$85 cm in men and $\geq$90 cm in women); however, this recommendation is based on one study, the validity of which has been called into question.[12,23,24] Epidemiological data collections in various populations are ongoing to create ethnic-specific criteria for abdominal obesity.

As mentioned, the description of "Syndrome X" in 1988 did not include obesity as part of the syndrome, as Reaven felt IR was the cause of this syndrome, not obesity. Studies investigating the relationship between obesity and IR have shown that although obese individuals are more likely to have IR

Not all obese individuals display the metabolic phenotype associated with increased risk of cardiovascular disease and type 2 diabetes mellitus. Greater degrees of insulin resistance have been associated with higher blood pressure, higher triglyceride levels, low high-density lipoprotein cholesterol, and increased rates of impaired fasting glucose.

and metabolic syndrome, not all obese individuals display the metabolic phenotype associated with increased risk of CVD and T2DM.[16,17] Greater degrees of IR have been associated with higher blood pressure, higher TG levels, low HDL, and increased rates of impaired fasting glucose.[25,26] Therefore, it appears that it is IR, which is frequently associated with obesity, that leads to the development of the metabolic syndrome.

PREVALENCE

The National Health and Nutrition Examination Survey 2003–2006 reported that 34% of the US population aged >20 years met the modified Adult Treatment Panel III (ATP III) criteria for the metabolic syndrome. These criteria are the same as those outlined in Table 1, except the cutoff for systolic blood pressure was ≥135 mm Hg and abdominal obesity was defined as a WC of ≥102 cm in men and ≥88 cm in women. The prevalence of the metabolic syndrome was similar in men and women, but increased with age and varied by ethnic group. The highest prevalence was found in non-Hispanic white men (37.2%) and Mexican American women (40.6%) (Table 2).[27] Using the ATP III criteria, in European studies the prevalence has been reported to be from 10% to 26%, rates in India have been reported to be from 7.9% to 46% in different populations, the prevalence in Iran has been reported as 24% in men and 42% in women, in Turkey 27% in men and 38.5% in women, and in Filipina American women aged 50–69 years living in the United States, the rate is reported as 34.3%.[28] Rates are lower in Japan, with a prevalence of 8.1% in men and 9.9% in women.[29] In individuals who have T2DM, the prevalence of the metabolic syndrome is reported to be 20%–94%, depending on the population studied.[30,31] The significance of diagnosing the metabolic syndrome in those with T2DM has been questioned, given that all patients with T2DM should have lifestyle and dietary intervention, glycemic control, blood pressure, and lipid management as part of their usual diabetes care and are already identified as having a higher risk of CVD.[31]

According to the National Health and Nutrition Examination Survey 2003–2006, 34% of the US population aged >20 years met the modified National Cholesterol Education Program Adult Treatment Panel III criteria for the metabolic syndrome.

Depending the population studied, the prevalence of the metabolic syndrome in individuals who have type 2 diabetes mellitus is 20%–94%.

Table 2. *Prevalence of Metabolic Syndrome by Selected Characteristics*[27].

Characteristic	%	OR	95% CI
Sex			
M	35.1	1.0	
F	32.6	0.89	(0.73–1.07)
M			
Age, yr			
20–39	20.3	1.0	
40–59	40.6	2.7	(1.96–3.73)
60+	51.5	4.18	(3.01–5.79)
Race and ethnicity*			
Non-Hispanic white	37.2	1.0	
Non-Hispanic black	25.3	0.54	(0.40–0.73)
Mexican American	33.2	0.78	(0.57–1.07)
BMI*			
Underweight-normal	6.8	1.0	
Overweight	29.8	6.17	(3.96–9.62)
Obese	65	31.92	(20.06–50.78)
F			
Age, yr			
20–39	15.6	1.0	
40–59	37.2	3.2	(2.32–4.43)
60+	54.4	6.44	(4.75–8.72)
Race and ethnicity*			
Non-Hispanic white	31.5	1.0	
Non-Hispanic black	38.8	1.44	(1.05–1.98)
Mexican American	40.6	1.55	(1.06–2.29)
BMI*			
Underweight-normal	9.3	1.0	
Overweight	33.1	5.48	(3.75–8.02)
Obese	56.1	17.14	(12.54–23.44)

Abbreviations: BMI, body-mass index; CI, confidence interval; F, female; M male; OR, odds ratio.

* Race and ethnicity and BMI data are adjusted for age.

PATHOGENESIS

Many factors contribute to the development of the metabolic syndrome. These include lifestyle characteristics such as diet and lack of physical activity, obesity, genetic predisposition, chronic inflammation, and the presence of cytokines, adipokines, elevated free fatty acids (FFAs), mitochondrial dysfunction, and alterations in insulin signaling. Uncommon genetic disorders in humans have helped to understand the physiological roles some of these factors play in the metabolic syndrome. Cases of defective insulin activity have been described with rare mutations in the insulin receptor gene (*INSR*) and downstream signaling molecules. The most-severe forms of *INSR* mutations lead to childhood or adolescent death. The less-severe forms cause significant IR, with β-cell compensation and hyperinsulinemia. In later adult life, diabetes may occur in these individuals due to β-cell dysfunction, akin to the development of T2DM following hyperinsulinemia and IR with the metabolic syndrome. Other metabolic abnormalities have not been well characterized in these individuals, as individuals with severe disease die young and less-severe forms may go undiagnosed.[32] However, inherited partial lipodystrophy syndromes have demonstrated the link between IR, visceral adiposity, dyslipidemia, T2DM, and hypertension. Various mutations have been found that lead to partial lipodystrophy syndromes. One of particular interest is the mutation of the gene for peroxisome proliferator-activated receptor-γ (PPAR-γ). PPAR-γ is a transcription factor that regulates the transcription of genes concerned with insulin sensitivity, adipocyte differentiation, and inflammation.[32,33] Compared with carriers of other mutations leading to partial lipodystrophy syndromes, carriers of mutations in the PPAR-γ gene have less adipose tissue loss, but more notable hyperinsulinemia and earlier onset of T2DM, with more abnormalities in adipokines and more severe hypertension, suggesting that the IR associated with PPAR-γ gene mutations may be more important than the adipocyte abnormalities in the development of the metabolic syndrome phenotype.[34]

Normal Insulin Action

In an insulin-sensitive individual, glucose stimulates insulin release from the pancreatic β cells. Insulin then decreases plasma glucose concentrations by suppressing hepatic glycogenolysis and gluconeogenesis, as well as by promoting glucose transport into muscle and adipose tissue. Glucose is taken up by muscle and adipose tissue by insulin-responsive glucose transporter 4 (GLUT4), which, under the influence of insulin, is mobilized from intracellular storage vesicles to the cell surface.[35] Exercise increases the expression of GLUT4, and its expression correlates with insulin sensitivity.[36] As well as its effects on glucose uptake, insulin prevents lipolysis in adipose tissue by dephosphorylating and thus inhibiting hormone-sensitive lipase. Hormone-sensitive lipase mediates the hydrolysis of TG in adipocytes with the release of FFAs and glycerol. Glycerol enters the circulation and is taken up by the liver and kidney, where it is converted to glucose by gluconeogenesis.

Insulin activates lipoprotein lipase in adipose tissue, allowing extraction of FFAs from very-low-density lipoprotein (VLDL). These FFAs are used for energy or are re-esterified into TG. In the liver, FFAs are taken up by the liver and processed into VLDL; however, insulin inhibits this process indirectly by decreasing the supply of FFAs to the liver through inhibition of adipose tissue lipogenesis.[37] In addition, insulin suppresses hepatic VLDL synthesis through phosphatidylinositol 3-kinase (PI3K)-mediated suppression of apolipoprotein B (apoB) production.[38] Acutely, insulin appears to inhibit fatty acid synthase, while with chronic hyperinsulinemia, insulin induces fatty-acid synthase activity.[39] In the setting of hyperinsulinemia and IR, there is increased flux of FFAs from adipose tissue to the liver with increased production of VLDL by the liver.[40]

Insulin is transported into the central nervous system (CNS) by insulin transporters in the blood-brain barrier. Low levels of insulin may also be synthesized locally in the brain, although this is the subject of debate.[41] Through insulin receptor signaling in the CNS, insulin decreases appetite.[42] Insulin receptors are found in the olfactory bulb, hypothalamus, hippocampus, amygdale, cerebral cortex, and cerebellum in rodents.[43–45] Binding of insulin to the insulin receptor leads to autophosphorylation of the insulin receptor and activation of signaling pathways, as described below, with phosphorylation of transcription factors that regulate the expression of the genes for neuropeptides, such as neuropeptide Y (NPY), pro-opiomelanocortin (POMC), and Agouti-related protein (AgRP). Insulin suppresses NPY synthesis in neurons of the arcuate nucleus of the hypothalamus, preventing the NPY stimulation of food ingestion. Insulin increases the synthesis of POMC in a second group of neurons of the arcuate nucleus. POMC is cleaved to form α-melanocyte stimulating hormone (α-MSH), which acts on melanocortin 3 receptors and melanocortin 4 receptors to decrease food intake.[46] Insulin signaling activates the adenosine triphosphate (ATP)-sensitive potassium (K_{ATP}) channels in the arcuate nucleus, decreases AgRP release, and through the vagus nerve, leads to decreased glucose production by the liver by decreasing gluconeogenic enzymes (phosphoenolpyruvate carboxykinase, glucose 6-phosphatase).[47,48] Additionally, CNS insulin signaling has been shown to increase adipose tissue lipoprotein lipase activity, adipose tissue mass, and adipocyte size.[49] In insulin-resistant states, there appears to be a decrease in cerebrospinal fluid insulin relative to serum insulin levels, leading to a relative CNS insulin deficit.[50] This is thought to be related to a decreased sensitivity of the insulin receptors in the blood-brain barrier to circulating insulin, in addition to resistance of the CNS structures to the action of insulin.[47,51–53]

Normal Insulin Signaling

Insulin can exert metabolic and mitogenic effects on cells through its binding to the insulin receptors and subsequent activation of signal transduction pathways. Insulin binding to the insulin receptors results in its autophosphorylation at sites on the intracellular β subunit. Autophosphorylation of the insulin receptors leads to the activation of tyrosine kinases and phosphorylation of other proteins. These proteins include insulin receptor substrates (IRS) and certain adaptor proteins that then act as docking sites for the phosphorylation of downstream effectors.[54] The pathways through which insulin mediates its metabolic activity are initiated by PI3K and the oncogene Cbl and its associated protein (CAP). Insulin transmits its growth effects through its action on the mitogen-activated protein kinase (MAPK) pathway.[55] Activation of the PI3K pathway ultimately leads to the activation of the protein kinase Akt and the phosphorylation of transcription factors, such as the FOXO1 member of the forkhead family of transcription factors. Phosphorylation of these transcription factors prevents their entry into the cell nucleus, thereby inhibiting their action. The FOXO1 transcription factor regulates the genes for hepatic gluconeogenesis; therefore insulin signaling decreases transcription of gluconeogenic genes.[56] FOXO1 also interacts with the promoter of another transcription factor, PPARγ. PPARγ regulates the GLUT4 gene transcription, FOXO1 inhibits the expression of PPARγ, whereas insulin signaling releases the suppression by FOXO1.[57] Signaling through a second pathway involves phosphorylation of the Cbl/CAP pathway and leads to actin-mediated mobilization of Glut4 to the cell surface, allowing for increased glucose uptake.[58] Activation of the MAPK pathway leads to phosphorylation of transcription factors involved in cell growth. These pathways interact with each other and are also regulated by other factors that may inhibit insulin signaling and contribute to IR. The elements involved in IR are outlined in Figure 1.

Factors Involved in Insulin Resistance

Free Fatty Acids

Human studies in nondiabetic, nonobese volunteers have shown that infusing lipid emulsions can cause IR and impaired glucose uptake in skeletal muscle.[59] Further, high FFA levels decrease Glut4 levels in human cardiac muscle.[60] Lowering FFA levels in obese nondiabetic individuals led to normalization of insulin-mediated glucose uptake by skeletal muscle. In individuals with obesity and T2DM, when FFA levels were lowered, insulin sensitivity improved.[61]

FFAs are absorbed from the intestine into the circulation in the form of chylomicrons. They also enter the circulation after lipolysis of adipose tissue. In obese individuals, FFAs are increased, despite elevated levels of insulin. This is thought to be due to increased adipose tissue lipolysis from a reduction in insulin's usual inhibition of hormone-sensitive lipase. The increase in FFAs then further impairs insulin signaling, leading to a greater increase in circulating FFA levels. In addition, there is reduced lipogenesis in the adipocytes due to decreased activity of transcription factors such as PPARγ.[62] As FFA levels increase, uptake into adipose tissue exceeds its oxidation and so metabolites involved in esterification accumulate. These include long-chain acyl coenzyme A (acyl CoA) and diacylglycerol (DAG). DAG also disrupts normal insulin signaling by serine phosphorylation of IRS, through activation of protein kinase C. Serine phosphorylation prevents tyrosine phosphorylation and thus impairs insulin signaling. Ceramide is another lipid molecule that accumulates in skeletal muscle of obese individuals. It activates the enzyme protein phosphatase 2A, leading to dephosphorylation of Akt, thereby blocking insulin signaling and inhibiting Glut4 translocation to the cell membrane, thus disrupting insulin-mediated glucose uptake into skeletal muscle.[63]

In the liver, FFAs inhibit the insulin-induced suppression of glycogenolysis and gluconeogenesis.[64] In animal studies, consuming a high-fat diet caused increased levels of FFAs and resulted in the loss of suppression of hepatic gluconeogenesis.[65] The increased supply of FFAs to the liver from adipocyte lipolysis enables the hepatic production and release of VLDL. Adipocytes also produce increased amounts of cholesterol ester transferase protein that decreases

> *Free fatty acids inhibit the insulin-induced suppression of glycogenolysis and gluconeogenesis.*

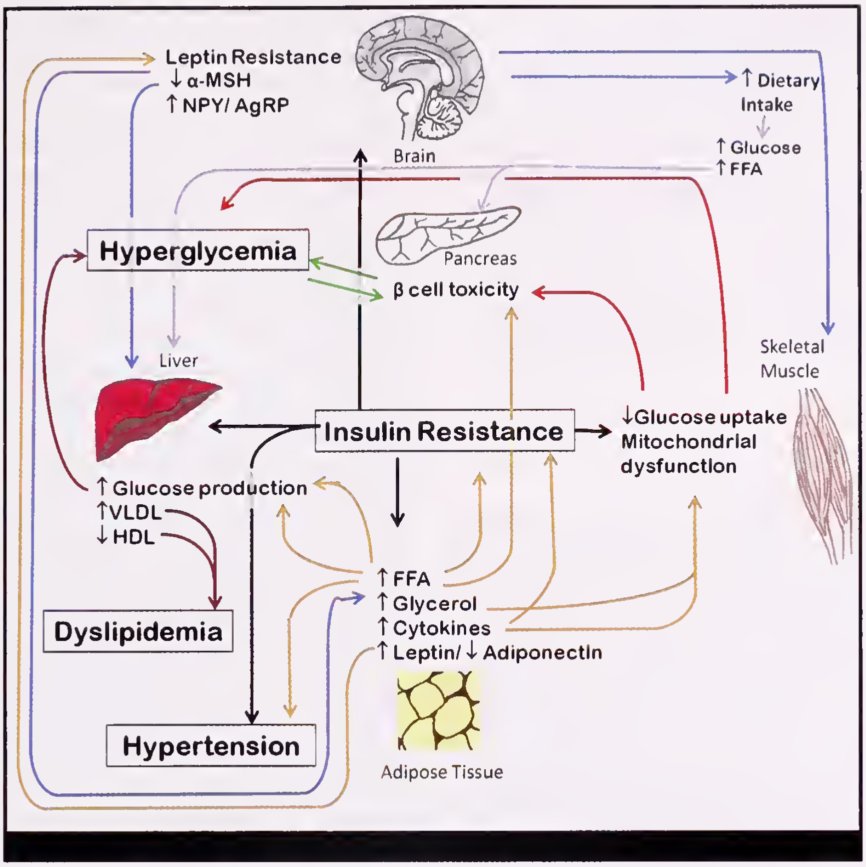

Leptin Resistance
↓α-MSH
↑NPY/ AgRP
Brain
↑Dietary Intake
↑Glucose
↑FFA
Hyperglycemia
Pancreas
β cell toxicity
Liver
Skeletal Muscle
Insulin Resistance
↓Glucose uptake
Mitochondrial dysfunction
↑Glucose production
↑VLDL
↓HDL
Dyslipidemia
↑FFA
↑Glycerol
↑Cytokines
↑Leptin/ ↓Adiponectin
Hypertension
Adipose Tissue

Fig 1. Insulin resistance in the development of the metabolic syndrome. Adipose tissue: leptin production increases due to leptin resistance, along with increased production of cytokines, diacyglycerol, and FFAs. Adiponectin levels decrease with increasing obesity. Adipose tissue production of cytokines, acyl Co A, diacylglycerol, and FFAs leads to insulin resistance, mitochondrial dysfunction, dyslipidemia, decreased glucose uptake by skeletal muscle, increased glucose production by the liver, and β-cell toxicity. Hyperglycemia ensues. The hyperglycemia can in turn cause further β-cell toxicity. Insulin resistance also leads to hypertension through its actions on the nitric oxide and endothelin-1 in the endothelium. Decreased insulin signaling in the brain results in increased dietary intake, as well as loss of suppression of hepatic gluconeogenesis and increased production of lipogenesis intermediated by adipose tissue. **Abbreviations:** acyl CoA, long-chain acyl coenzyme A; AgRP, Agouti-related protein; α-MSH, α-melanocyte stimulating hormone; FFA, free fatty acids; HDL, high-density lipoprotein cholesterol; NPY, neuropeptide Y; VLDL, very-low-density lipoprotein cholesterol.

HDL levels by facilitating the transfer of cholesteryl esters from HDL to VLDL.[66] This contributes to low HDL and increased TG levels. In addition, increased hepatic lipase degrades HDL and the liver exchanges VLDL for HDL.[67] Finally, there is evidence that the HDL level in insulin-resistant individuals with and without hypertriglyceridemia is largely determined by the fractional catabolic rate of apolipoprotein A-I and A-II (the major lipoproteins that constitute HDL). It has been suggested that individuals with low HDL levels, such as those with the metabolic syndrome, have more loosely bound apoA-1 that is more readily catabolized by the kidneys.[68,69]

FFAs also contribute to the development of nonalcoholic fatty liver disease in insulin-resistant individuals. When FFAs are released from visceral adipose tissue into the portal circulation and from subcutaneous adipose tissue into the peripheral circulation, they may be taken up by hepatocytes and bound to CoA. Fatty acyl CoA can then form hepatic TG or interfere with insulin signaling and induce inflammation in hepatocytes. In insulin-resistant states, insulin can activate the transcription factor SREBP-1, increasing the transcription of genes involved in hepatic lipogenesis. Glucose activates another transcription factor, carbohydrate response element binding protein (ChREBP). ChREBP stimulates the glycolysis of glucose into pyruvate, which forms acetyl CoA and subsequently malonyl CoA, which is involved in FFA synthesis. Both ChREBP and SREBP transcription factors are increased in animal models of fatty liver. Additionally, de novo lipogenesis is increased, possibly through activation of the transcription factor FOXA2. In mouse studies, IR is associated with impaired Foxo1 signaling, but Foxa2 phosphorylation is preserved. Together, these alterations in glucose and lipid metabolism can lead to hepatic lipogenesis and the development of nonalcoholic fatty liver disease.[62,70,71]

Inflammatory Cytokines

FFAs also increase inflammation by signaling through the receptor that normally interacts with the lipopolysaccharide membrane of gram-negative bacteria (Toll-like receptor 4). This interaction causes activation of the nuclear factor κB transcription factor and increased synthesis of cytokines and chemokines, including tumor necrosis factor α (TNFα); interleukins (IL) IL-1 β and IL-6; and monocyte chemoattractant protein-1 (MCP-1). This leads to a state of chronic inflammation in obese individuals.

The inflammatory cytokines IL-6 and TNFα are produced by adipose tissue macrophages. Cytokines can interfere with insulin action by serine phosphorylation of IRS and prevent downstream signaling through the PI3K pathway.[72] They trigger the production of proinflammatory cytokines that also cause serine phosphorylation of IRS and impair insulin signaling. These proinflammatory cytokines up-regulate proinflammatory genes and thus propagate the inflammatory signaling. TNFα also decreases endothelial nitric oxide synthase (eNOS), leading to decreased expression of mitochondrial oxidative phosphorylation genes, increasing oxidative cellular stress and the accumulation of reactive oxygen species (ROS).

Mitochondrial Dysfunction

In the cell, the mitochondria are involved in energy production in the form of ATP, through oxidative phosphorylation and the transfer of electrons through the respiratory chain. There is a link between mitochondrial dysfunction and IR, before the development of obesity and T2DM. Insulin-resistant, physically inactive, obese individuals and those with T2DM have smaller and fewer mitochondria in skeletal muscle. In addition, studies in humans have shown a down-regulation of genes involved in mitochondrial oxidative phosphorylation. Therefore, insulin-resistant individuals may have a lower number of mitochondria, with impaired function. The exact mechanisms by which IR leads to mitochondrial dysfunction are unclear, but there appears to be decreased expression of the transcription factor coregulator of nuclear receptors coactivator 1α (PGC-1α) in insulin-resistant individuals.[73,74]

Mitochondrial dysfunction leads to decreased lipid oxidation, with the accumulation of lipid metabolites DAG, ceramide, and acyl CoA. These in turn cause impaired insulin signaling, as described. It also leads to increased oxidative stress and increased ROS formation, and a vicious cycle ensues as these elements contribute to further mitochondrial dysfunction. Impairment of mitochondrial activity may lead to decreased ATP formation and the perception of an energy deficit by the brain, causing appetite stimulation.[73]

The production of ATP by the mitochondria is driven by a proton gradient across the inner mitochondrial membrane. Mitochondria may generate heat by the "proton leak" as protons leak from the intermembrane space across the inner mitochondrial membrane into the matrix. There are 3 uncoupling protein that play a major role in this process. Mitochondrial uncoupling results in elevation of long-chain acyl CoA and the formation of DAG, which in turn activates protein kinase C, which interferes with normal insulin signaling.[75]

Adipokines

Leptin is a protein produced by adipocytes that controls appetite and energy expenditure. Absence of leptin leads to extreme obesity and IR. However, most obese individuals do not have a leptin deficiency; instead, they have increased levels of leptin, but are resistant to its appetite-suppressing effects.[76] Elevated TG levels decrease leptin transport into the CNS, which, along with impaired leptin signaling, may lead to leptin resistance.[44,48] The leptin receptors are mostly located in the brain, adipocytes, and endothelial cells. Many of the metabolic effects of leptin are via the hypothalamus, through its regulation of NPY and α-MSH, in a manner similar to insulin.[48,76] Animal studies have shown that leptin promotes lipid oxidation and protein synthesis. In studies on rat adipocytes, leptin reduces insulin's lipogenic effects, potentially due to its inhibition of insulin binding to the adipocyte.[77] Leptin also depletes triglyceride content of adipose tissue without increasing circulating FFAs, due to increased mitochondrial oxidation. In this way it also opposes the actions of insulin.[76–78] Leptin does not seem to interfere with insulin actions on glycogen synthesis or glycolysis. As well as its effects on lipid, protein, and glucose metabolism, leptin also stimulates T-cell proliferation, increases the production of inflammatory cytokines, and regulates the reproductive system.[79] Leptin resistance leads to loss of leptin's appetite suppression and increased adiposity.

Adiponectin is another protein produced by adipocytes. It has anti-inflammatory and antiatherogenic properties and regulates food intake.[80–83] Adiponectin levels are absent in mice without adipose tissue; however, IR and increased intra-abdominal fat is also associated with decreased levels of adiponectin.[84,85] Binding of adiponectin to its receptors (AdipoR1 and AdipoR2) leads to the activation of several signaling pathways involved in the regulation of glucose, protein, and fatty-acid metabolism. Phosphorylation of AMP-activated protein kinase (AMPK) as a result of adiponectin signaling leads to increased glucose uptake in the muscle and decreased hepatic gluconeogenesis. Adiponectin deficiency is also linked to increased atherogenesis, elevated levels of small dense LDL and TG levels, along with increased inflammatory cytokines.[77,85] The exact mechanisms through which IR leads to low adiponectin levels and the involvement of adiponectin in the development of the metabolic syndrome remain unclear.[86]

Resistin is another adipokine that is increased in animal models of obesity. It has been shown to lead to impaired insulin action and glucose intolerance. It appears to cause β-cell apoptosis in rat insulinoma cells and so has been postulated to lead to β-cell dysfunction in T2DM.[87] In addition, resistin activates the suppressor of cytokine signaling proteins that inhibit phosphorylation of the insulin receptor and downstream insulin signaling proteins, thus leading to impaired insulin signaling.[88] It also impairs glucose uptake by skeletal muscle and the liver and increases hepatic gluconeogenesis.[89,90] Other adipokines that have been identified include apelin, visfatin, and retinol binding protein-4. Some studies have found positive correlations between these adipokines and IR, although their exact roles and mechanisms remain to be determined.[91]

METABOLIC EFFECTS OF INSULIN RESISTANCE

Together, abnormalities in insulin sensitivity, FFA levels, cytokine production, mitochondrial function, and levels of adipokines contribute to the development of hypertension, dyslipidemia, and hyperglycemia and comprise the metabolic syndrome (Figure 1).

> *Abnormalities in insulin sensitivity, free fatty acid levels, cytokine production, mitochondrial function, and levels of adipokines contribute to the development of hypertension, dyslipidemia, and hyperglycemia and comprise the metabolic syndrome.*

Hypertension and Insulin Resistance

Hypertension is present in 15% of males and females with a BMI of $\leq$25 kg/m^2 and increases to 42% of males and 38% of females with a BMI of $>$30 kg/m^2.[92,93] Studies have shown that IR or hyperinsulinemia is present in the majority of hypertensive patients, suggesting a link between obesity, glucose intolerance, and hypertension. Insulin is a stimulator of the vasodilator nitric oxide (NO) through insulin-mediated PI3K signaling and phosphorylation of eNOS. Under normal circumstances, vasodilatation occurs in the skeletal-muscle vasculature in response to insulin release after eating and promotes glucose disposal.[94] In opposition to NO, endothelin-1 (ET-1) is a potent vasoconstrictor, which is stimulated by insulin activity through the MAPK pathway. NO is an inhibitor of ET-1. However, it has been proposed that in insulin-resistant states, insulin signaling through the PI3K pathway is impaired, leading to decreased NO. In addition, increased cytokines, such as IL-6, along with low adiponectin levels and leptin resistance, cause decreased NO and increased action of ET-1, with the subsequent development of hypertension. FFA can also induce hypertension by increased production of ROS, which increase oxidative stress and decrease NO.[93,94] Hyperinsulinemia may also lead to increased peripheral vascular resistance due to sympathetic overactivity, volume expansion from its antinatriuretic effects, and increased antiotensinogen II.

Dyslipidemia and Insulin Resistance

The dyslipidemic profile of the metabolic syndrome involves low HDL and elevated TG. In a meta-analysis of prospective studies, it was reported that those with TG levels in the top third of the population had a relative risk of coronary disease of 1.7 (95% confidence interval: 1.6–1.9) compared with those with TG levels in the bottom third.[95] The onset of IR has a profound effect on lipid profiles.[96] IR results in the increased production of FFA from adipocytes through loss of its inhibition of hormone-sensitive lipase. Additionally, endothelial lipoprotein lipase function is impaired, both events leading to the increase in circulating FFAs. The increased influx of FFAs to the liver, and insulin stimulation of hepatic lipogenesis, lead to increased hepatic TG production in the form of VLDL and steatosis as the TG are stored in the liver.[35,67] Adipocyte production of cholesterol ester transferase protein allows for the transfer of cholesteryl esters from HDL to VLDL. There is increased clearance of HDL by the kidneys and the liver takes up HDL and produces VLDL, therefore leading to the low HDL and elevated TG levels seen with the metabolic syndrome.[66,67,69]

> *The increased influx of FFAs to the liver, and insulin stimulation of hepatic lipogenesis, lead to increased hepatic triglyceride production in the form of very-low-density lipoprotein and steatosis as the triglycerides are stored in the liver.*

Hyperglycemia and Insulin Resistance

IR is a key element in the development of hyperglycemia and T2DM. In normal individuals, ingestion of nutrients results in the secretion of insulin from the pancreatic β cells. Specific triggers for insulin release include glucose, amino acids, FFAs, and gastrointestinal hormones, such as glucagon like peptide-1. There is a rapid first-phase insulin response that occurs immediately after eating, peaks at 10 minutes, and disappears after approximately 20 minutes. Its effect is to inhibit hepatic glucose production and promote glucose uptake. This is followed by a second-phase insulin response, which begins at 15–20 minutes and peaks during the next 20–40 minutes.[97]

In IR and T2DM, the initial β cell abnormality is loss of the first-phase insulin response, followed by an exaggerated second-phase response. This ultimately leads to hyperinsulinemia. Chronic nutrient overload leads to β-cell toxicity and eventually β-cell failure.

This can be seen in animal models, where chronic hyperglycemia leads to increased basal insulin levels, but loss of β-cell response to glucose stimulation. In the presence of hyperglycemia, FFAs can also cause β-cell dysfunction, by inducing enzymes of β-oxidation, leading to increased acetyl CoA levels, activation of pyruvate carboxylase and pyruvate cycling in the islet cells. This also causes elevated basal insulin levels and impaired glucose stimulated insulin release. The toxic effects of glucose and lipids are also seen in human β cells, and are referred to as glucotoxicity and lipotoxicity, respectively.[98,99]

> *Chronic nutrient overload leads to β-cell toxicity and eventually β-cell failure. Human β cells also show the toxic effects of glucose and lipids, which are referred to as glucotoxicity and lipotoxicity, respectively.*

Increased demand for insulin leads to endoplasmic reticulum stress and cell death. Amylin, which is hypersecreted with insulin from the β cell, can form amyloid fibrils that can accumulate in β-cells and also induce β-cell dysfunction. In normal circumstances, in insulin-sensitive individuals, there is a feedback loop between the β cells and the liver, skeletal muscle, and adipose tissue. To maintain euglycemia, the β cells have to increase or decrease secretion in response to signals from these tissues. Failure of the feedback loop leads to the development of impaired glucose tolerance and T2DM.[98,99]

CONCLUSION

The presence of the metabolic syndrome should be considered in the evaluation of overweight and obese patients, as it may identify those at greatest risk of developing T2DM and CVD in the future. The diagnostic criteria have now been harmonized and simplified from the previous versions to provide a practical and useful set of criteria that can be used in the physician's office and allows agreement for comparison amongst epidemiological studies. Although the diagnosis may be less complicated, the mechanisms by which the metabolic syndrome develops are the subject of ongoing research and are becoming more complex, with interplay between insulin signaling, inflammation, mitochondrial function, adipokines, FFAs, lipid production, glucose uptake, genetics, lifestyle, and diet. While obesity may exaggerate the metabolic syndrome phenotype, it appears that IR is the key element that can be present in both obese and lean individuals and leads to the derangements in lipids, blood pressure, and glycemic control that characterize the metabolic syndrome.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Haller H. Epidemiology and associated risk factors of hyperlipoproteinemia [in German]. *Z Gesamte Inn Med* 1977; 32: 124–128.
2. Chadwick J, Mann WN. *Medical Works of Hippocrates.* Boston, MA: Blackwell Scientific Publications; 1950; 154.
3. Erkelens DW, deBruin TW, Cabezas MC. Tulp syndrome. *Lancet* 1993; 342: 1536–1537.
4. Enzi G, Busetto L, Inelmen EM, et al. Historical perspective: visceral obesity and related comorbidity in Joannes Baptista Morgagni's 'De Sedibus et Causis Morborum per Anatomen Indagata'. *Int J Obes Relat Metab Disord* 2003; 27: 534–535.
5. Nilsson S. Research contributions of Eskil Kylin [in Swedish]. *Sven Med Tidskr* 2001; 5: 15–28.
6. Reaven GM. Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes* 1988; 37: 1595–1607.
7. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998; 15: 539–553.
8. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III). *J Am Med Assoc* 2001; 285: 2486–2497.
9. Einhorn D, Reaven GM, Cobin RH, et al. American College of Endocrinology position statement on the insulin resistance syndrome. *Endocr Pract* 2003; 9: 237–252.
10. Alberti KG, Zimmet P, Shaw J. The metabolic syndrome–a new worldwide definition. *Lancet* 2005; 366: 1059–1062.
11. Grundy S, Cleeman J, Daniels S, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute scientific statement. *Circulation* 2005; 112: 2735–2752.
12. Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009; 120: 1640–1645.
13. Ford ES. Risks for all-cause mortality, cardiovascular disease, and diabetes associated with the metabolic syndrome: a summary of the evidence. *Diabetes Care* 2005; 28: 1769–1778.
14. Gale E. Should we dump the metabolic syndrome? Yes. *BMJ* 2008; 336: 640.
15. Kahn R. The metabolic syndrome (Emperor) wears no clothes. *Diabetes Care* 2006; 29: 1693–1696.
16. Brochu M, Tchernof A, Dionne IJ, et al. What are the physical characteristics associated with a normal metabolic profile despite a high level of obesity in postmenopausal women? *J Clin Endocrinol Metab* 2001; 86: 1020–1025.
17. Stefan N, Kantartzis K, Machann J, et al. Identification and characterization of metabolically benign obesity in humans. *Arch Intern Med* 2008; 168: 1609–1616.
18. Grundy SM. Does the metabolic syndrome exist? *Diabetes Care* 2006; 29: 1689–1692.
19. Kuk JL, Ardern CI. Are metabolically normal but obese individuals at lower risk for all-cause mortality? *Diabetes Care* 2009; 32: 2297–2299.
20. Yusuf S, Hawken S, Ounpuu S, et al. Obesity and the risk of myocardial infarction in 27,000 participants from 52 countries: a case-control study. *Lancet* 2005; 366: 1640–1649.
21. Ohlson LO, Larsson B, Svärdsudd K, et al. The influence of body fat distribution on the incidence of diabetes mellitus: 13.5 years of follow-up of the participants in the study of men born in 1913. *Diabetes* 1985; 34: 1055–1058.
22. Lear SA, Humphries SK, Birmingham CL. The use of BMI and waist circumference as surrogates of body fat differs by ethnicity. *Obesity* 2007; 15: 2817–2824.
23. Oka R, Kobayashi J, Yagi K, et al. Reassessment of the cutoff values of waist circumference and visceral fat area for identifying Japanese subjects at risk for the metabolic syndrome. *Diabetes Res Clin Pract* 2008; 79: 474–481.
24. Examination Committee of Criteria for 'Obesity Disease' in Japan, Japan Society for the Study of Obesity. New criteria for 'obesity disease' in Japan. *Circ J* 2002; 66: 987–992.
25. Farin HM, Abbasi F, Reaven GM. Comparison of body mass index versus waist circumference with metabolic changes that increase the risk of cardiovascular disease in insulin resistant individuals. *Am J Cardiol* 2006; 98: 1053–1056.
26. Reaven GM. Insulin resistance: the link between obesity and cardiovascular disease. *Endocrinol Metab Clin North Am* 2008; 37: 581–601.
27. Ervin RB. Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003–2006. *Natl Health Stat Report* 2009; 13: 1–7.
28. Cameron AJ, Shaw JE, Zimmet PZ. The metabolic syndrome: prevalence in worldwide populations. *Endocrinol Metab Clin North Am* 2004; 33: 351–375.
29. Lee CM, Huxley RR, Woodward M, et al. Comparisons of metabolic syndrome definitions in four populations of the Asia-Pacific region. *Metab Syndr Relat Disord* 2008; 6: 37–46.
30. Sone H, Tanaka S, Ishibashi S, et al. The new worldwide definition of metabolic syndrome is not a better diagnostic predictor of cardiovascular disease in Japanese diabetic patients than the existing definitions. *Diabetes Care* 2006; 29: 145–147.

31. Song SH, Hardisty CA. Diagnosing metabolic syndrome in type 2 diabetes: does it matter? *QJM* 2008; 101: 487–491.
32. Hegele RA. Monogenic forms of insulin resistance: apertures that expose the common metabolic syndrome. *Trends Endocrinol Metab* 2003; 14: 371–377.
33. O'Rahilly S. Human obesity and insulin resistance: lessons from experiments of nature. *Biochem Soc Trans* 2007; 35: 33–36.
34. Hegele RA. Lessons from human mutations in PPARgamma. *Int J Obes (Lond)* 2005; 29(suppl 1): S31–S35.
35. Karnieli E, Zarnowski MJ, Hissin PJ, et al. Insulin-stimulated translocation of glucose transport systems in the isolated rat adipose cell: time course, reversal, insulin concentration dependency and relationship to glucose transport activity. *J Biol Chem* 1981; 256: 4772–4777.
36. Karnieli E, Armoni M. Transcriptional regulation of the insulin-responsive glucose transporter GLUT4 gene: from physiology to pathology. *Am J Physiol Endocrinol Metab* 2008; 295: E38–E45.
37. Stefan N, Kantartzis K, Häring HU. Causes and metabolic consequences of fatty liver. *Endocr Rev* 2008; 29: 939–960.
38. Chirieac DV, Davidson NO, Sparks CE, Sparks JD. PI3-kinase activity modulates apo B available for hepatic VLDL production in apobec-1$^{-/-}$ mice. *Am J Physiol Gastrointest Liver Physiol* 2006; 291: G382–G388.
39. Najjar SM, Yang Y, Fernström MA, et al. Insulin acutely decreases hepatic fatty acid synthase activity. *Cell Metab* 2005; 2: 43–53.
40. Zammit VA. Insulin stimulation of hepatic triacylglycerol secretion in the insulin-replete state: implications for the etiology of peripheral insulin resistance. *Ann N Y Acad Sci* 2002; 967: 52–65.
41. Schechter R, Beju D, Gaffney T, et al. Preproinsulin I and II mRNAs and insulin electron microscopic immunoreactions are present within the rat fetal nervous system. *Brain Res* 1996; 736: 16–27.
42. Air EL, Benoit SC, Blake Smith KA, et al. Acute third ventricular administration of insulin decreases food intake in two paradigms. *Pharmacol Biochem Behav* 2002; 72: 423–429.
43. Marks JL, Porte D Jr, Stahl WL, Baskin DG. Localization of insulin receptor mRNA in rat brain by in situ hybridization. *Endocrinology* 1990; 127: 3234–3236.
44. Isaganaitis E, Lustig RH. Fast food, central nervous system insulin resistance, and obesity. *Arterioscler Thromb Vasc Biol* 2005; 25: 2451–2462.
45. Banks WA. The source of cerebral insulin. *Eur J Pharm* 2004; 490: 5–12.
46. Woods SA, D'Alessio DA. Central control of body weight. *JCEM* 2008; 93: S37–S50.
47. Pocai A, Lam TK, Gutierrez-Juarez R, et al. Hypothalamic K(ATP) channels control hepatic glucose production. *Nature* 2005; 434: 1026–1031.
48. Pagotto U. Where does insulin resistance start? The brain. *Diabetes Care* 2009; 32(suppl 2): S174–S177.
49. Koch L, Wunderlich FT, Seibler J, et al. Central insulin action regulates peripheral glucose and fat metabolism in mice. *J Clin Invest* 2008; 118: 2132–2147.
50. Kern W, Benedict C, Schultes B, et al. Low cerebrospinal fluid insulin levels in obese humans. *Diabetologia* 2006; 49: 2790–2792.
51. Schwartz MW, Figlewicz DP, Kahn SE, et al. Insulin binding to brain capillaries is reduced in genetically obese, hyperinsulinemic Zucker rats. *Peptides* 1990; 11: 467–472.
52. Woods SC, Seeley RJ, Baskin DG, Schwartz MW. Insulin and the blood-brain barrier. *Curr Pharm Res* 2003; 9: 795–800.
53. Hallschmid M, Benedict C, Schultes B, et al. Obese men respond to cognitive but not to catabolic brain insulin signaling. *Int J Obes* 2008; 32: 275–282.
54. Paz K, Hemi R, LeRoith D, et al. A molecular basis for insulin resistance: elevated serine/threonine phosphorylation of IRS-1 and IRS-2 inhibits their binding to the juxtamembrane region of the insulin receptor and impairs their ability to undergo insulin-induced tyrosine phosphorylation. *J Biol Chem* 1997; 72: 29911–29918.
55. Taniguchi CM, Emanuelli B, Kahn CR. Critical nodes in signalling pathways: insights into insulin action. *Nat Rev Mol Cell Biol* 2006; 7: 85–96.
56. Puigserver P, Rhee J, Donovan J, et al. Insulin-regulated hepatic gluconeogenesis through FOXO1-PGC-1alpha interaction. *Nature* 2003; 423: 550–555.
57. Armoni M, Harel C, Karnieli E. Transcriptional regulation of the GLUT4 gene: from PPAR-γ and FOXO1 to FFA and inflammation. *Trends Endocrinol Metab* 2007; 18: 100–107.
58. Saltiel AR, Kahn CR. Insulin signaling and the regulation of glucose and lipid metabolism. *Nature* 2001; 414: 799–806.
59. Boden G, Jadali F, While J, et al. Effects of fat on insulin-stimulated carbohydrate metabolism in normal men. *J Clin Invest* 1991; 88: 960–966.
60. Armoni M, Harel C, Bar-Yoseph F, et al. Free fatty acids repress the GLUT4 gene expression in cardiac muscle via novel response elements. *J Biol Chem* 2005; 280: 34786–34795.
61. Santomauro AT, Boden G, Silva ME, et al. Overnight lowering of free fatty acids with Acipimox improves insulin resistance and glucose tolerance in obese diabetic and nondiabetic subjects. *Diabetes* 1999; 48: 1836–1841.
62. Guilherme A, Virbasius JV, Puri V, Czach MP. Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. *Nat Rev Mol Cell Biol* 2008; 9: 367–377.
63. Bonen A, Parolin ML, Steinberg GR, et al. Triacylglycerol accumulation in human obesity and type 2 diabetes is associated with increased rates of skeletal muscle fatty acid transport and increased sarcolemmal FAT/CD36. *FASEB J* 2004; 18: 1144–1146.
64. Boden G, Chen X, Ruiz J, et al. Mechanisms of fatty acid-induced inhibition of glucose uptake. *J Clin Invest* 1994; 93: 2438–2446.
65. Samuel VT, Liu ZX, Qu X, et al. Mechanism of hepatic insulin resistance in non-alcoholic fatty liver disease. *J Biol Chem* 2004; 279: 32345–32353.
66. Dullaart RP, Sluiter WJ, Dikkeschei LD, et al. Effect of adiposity on plasma lipid transfer protein activities: a possible link between insulin resistance and high density lipoprotein metabolism. *Eur J Clin Invest* 1994; 24: 188–194.
67. Boden G. Obesity and free fatty acids. *Endocrinol Metab Clin North Am* 2008; 37: 635–646.
68. Brinton EA, Eisenberg S, Breslow JL. Human HDL cholesterol levels are determined by apoA-I fractional

catabolic rate, which correlates inversely with estimates of HDL particle size. Effects of gender, hepatic and lipoprotein lipases, triglyceride and insulin levels, and body fat distribution. *Arterioscler Thromb* 1994; 14: 707–720.

69. Pietzsch J, Julius U, Nitzsche S, Hanefeld M. In vivo evidence for increased apolipoprotein A-I catabolism in subjects with impaired glucose tolerance. *Diabetes* 1998; 47: 1928–1934.
70. Hwang JH, Stein DT, Barzilai N, et al. Increased intrahepatic triglyceride is associated with peripheral insulin resistance: in vivo MR imaging and spectroscopy studies. *Am J Physiol Endocrinol Metab* 2007; 293: E1663–E1669.
71. Korenblat KM, Fabbrini E, Mohammed BS, Klein S. Liver, muscle and adipose tissue insulin action is directly related to intrahepatic triglyceride content in obese subjects. *Gastroenterology* 2008; 134: 1369–1375.
72. Shi H, Kokoeva MV, Inouye K, et al. TLR4 links innate immunity and fatty acid-induced insulin resistance. *J Clin Invest* 2006; 116: 3015–3025.
73. Kim JA, Wei Y, Sowers JR. Role of mitochondrial dysfunction in insulin resistance. *Circ Res* 2008; 102: 401–414.
74. Højlund K, Mogensen M, Sahlin K, Beck-Nielsen H. Mitochondrial dysfunction in type 2 diabetes and obesity. *Endocrinol Metab Clin North Am* 2008; 37: 713–731.
75. Erion DM, Schulman GI. Diacylglycerol-mediated insulin resistance. *Nat Med* 2010; 16: 400–402.
76. Buettner C, Muse ED, Cheng A, et al. Leptin controls adipose tissue lipogenesis via central, STAT3-independent mechanisms. *Nat Med* 2008; 14: 667–675.
77. Lago F, Gömez R, Gömez-Reino JJ, et al. Adipokines as novel modulators of lipid metabolism. *Trends Biochem Sci* 2009; 34: 500–510.
78. Shimabukuro M, Koyama K, Chen G, et al. Direct antidiabetic effect of leptin through triglyceride depletion of tissues. *Proc Natl Acad Sci U S A* 1997; 94: 4637–4641.
79. Rousli N, Kerin PA. Adipocytokines and the metabolic complications of obesity. *J Clin Endocrinol Metab* 2008; 93(suppl 1): S64–S73.
80. Scherer PE, Williams S, Fogliano M, et al. A novel serum protein similar to C1q, produced exclusively in adipocytes. *J Biol Chem* 1995; 270: 26746–26749.
81. Hu E, Liang P, Spiegelman BM. AdipoQ is a novel adipose-specific gene dysregulated in obesity. *J Biol Chem* 1996; 271: 10697–10703.
82. Mantzoros CS, Li T, Manson JE, et al. Circulating adiponectin levels are associated with better glycemic control, more favorable lipid profile, and reduced inflammation in women with type 2 diabetes. *J Clin Endocrinol Metab* 2005; 90: 4542–4548.
83. Kubota N, Yano W, Kubota T, et al. Adiponectin stimulates AMP activated protein kinase in the hypothalamus and increases food intake. *Cell Metab* 2007; 6: 55–68.
84. Arita Y, Kihara S, Ouchi N, et al. Paradoxical decrease of an adipose-specific protein, adiponectin, in obesity. *Biochem Biophys Res Commun* 1999; 257: 79–83.
85. Cnop M, Havel PJ, Utzschneider KM, et al. Relationship of adiponectin to body fat distribution, insulin sensitivity and plasma lipoproteins: evidence for independent roles of age and sex. *Diabetologia* 2003; 46: 459–469.
86. Cook JR, Semple RK. Hypoadiponectinemia–cause or consequence of human "insulin resistance"? *J Clin Endocrinol Metab* 2010; 95: 1544–1554.
87. Gao CL, Zhao DY, Zhang CM, et al. Resistin induces rat insulinoma cell RINm5F apoptosis. *Mol Biol Rep* 2009; 36: 1703–1708.
88. Steppan CM, Wang J, Whiteman EL, et al. Activation of SOCS-3 by resistin. *Mol Cell Biol* 2005; 25: 1569–1575.
89. Banerjee RR, Rangwala SM, Shapiro JS, et al. Regulation of fasted blood glucose by resistin. *Science* 2004; 303: 1195–1198.
90. Yang Y, Xiao M, Mao Y, et al. Resistin and insulin resistance in hepatocytes: resistin disturbs glycogen metabolism at the protein level. *Biomed Pharmacother* 2009; 63: 366–374.
91. Maury E, Brichard SM. Adipokine dysregulation, adipose tissue inflammation and metabolic syndrome. *Mol Cell Endocrinol* 2010; 314: 1–16.
92. Brown CD, Higgins M, Donato KA, et al. Body mass index and the prevalence of hypertension and dyslipidemia. *Obes Res* 2000; 8: 605–619.
93. Poirier P, Giles TD, Bray GA, et al. Obesity and cardiovascular disease: pathophysiology, evaluation and effect of weight loss: an update of the 1997 American Heart Association scientific statement on obesity and heart disease from the Obesity Committee of the Council on Nutrition, Physical Activity and Metabolism. *Circulation* 2006; 113: 898–918.
94. Muniyappa R, Iantorro M, Quon MJ. An integrated view of insulin resistance and endothelial dysfunction. *Endocrinol Metab Clin North Am* 2008; 37: 685–711.
95. Sarwar N, Danesh J, Gudny E, et al. Triglycerides and the risk of coronary heart disease. *Circulation* 2007; 115: 450–458.
96. Ginsberg HN. Insulin resistance and cardiovascular disease. *J Clin Invest* 2000; 106: 453–458.
97. LeRoith D. Beta-cell dysfunction and insulin resistance in type 2 diabetes: role of metabolic and genetic abnormalities. *Am J Med* 2002; 113(suppl 6A): 3S–11S.
98. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* 2006; 444: 840–846.
99. Muoio DM, Newgard CB. Mechanisms of disease: molecular and metabolic mechanisms of insulin resistance and beta-cell failure in type 2 diabetes. *Nat Rev Mol Cell Biol* 2008; 9: 193–205.

Mount Sinai Journal of Medicine 77:524–532, 2010

Therapeutic Options for the Prevention of Type 2 Diabetes Mellitus in the Metabolic Syndrome

Ivana Lukacova-Zib, MD and Geetha Gopalakrishnan, MD

Warren Alpert Medical School of Brown University, Providence, RI

OUTLINE

Lifestyle Interventions
- **Weight Loss**
- **Diet and Exercise**
- **Smoking, Alcohol Intake, and Coffee Consumption**

Therapeutic Interventions
- **Antidiabetic Agents**
- **Treatment for Obesity**
- **Antilipid and Antihypertensive Drugs**

Conclusion

ABSTRACT

The increasing prevalence of metabolic syndrome has created an impetus to prevent complications such as diabetes. There are no defined treatments for metabolic syndrome apart from addressing each of its components such as obesity, hypertension, dyslipidemia, and hyperglycemia. To evaluate diabetes-prevention strategies, we performed a Medline search and identified evidence suggesting that lifestyle intervention and certain medications can prevent diabetes. In addition, there are several agents on the horizon that seem promising. The mechanism responsible for diabetes prevention relates to improved insulin resistance as a result of weight loss or other processes. Lifestyle interventions and medications such as metformin are effective and safe in long-term prevention studies. However, the safety and efficacy of newer agents need clarification. ***Mt Sinai J Med 77:524–532, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: insulin resistance, metabolic syndrome, prediabetes, prevention of diabetes.

Address Correspondence to:

Geetha Gopalakrishnan
Warren Alpert Medical
School of Brown University
Hallett Center for Diabetes and
Endocrinology
Providence, RI
Email: ggopala@lifespan.org

Metabolic syndrome describes a set of risk factors that predispose individuals to cardiovascular disease and diabetes. The first formalized definition of the metabolic syndrome was proposed in 1998 by the World Health Organization.[1] Since then, modifications of the diagnostic criteria have been proposed and debated. Although dispute continues regarding the exact definition, a unified definition of metabolic syndrome was proposed in 2009 by several organizations including the International Diabetes Federation Task Force on Epidemiology and Prevention; the National Heart, Lung, and Blood Institute; the American Heart Association; the World Heart Federation; the International Atherosclerosis Society; and the International Association for the Study of Obesity.[2]

The current criteria for the diagnosis of metabolic syndrome includes 3 of the 5 following characteristics: elevated waist circumference (population-specific and country-specific definitions), elevated triglycerides (TG $\geq$150 mg/dL), reduced high-density lipoprotein cholesterol (HDL-C $\leq$40 mg/dL in males and $\leq$50 mg/dL in females), elevated blood pressure ($\geq$130/85 mm Hg), and elevated fasting glucose ($\geq$100 mg/dL). Although the rate of progression to diabetes can vary among different populations, several of the characteristics described above have been linked to diabetes, and the diagnosis of metabolic syndrome appears to

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20204

© 2010 Mount Sinai School of Medicine

Table 1. *Lifestyle Modifications and Therapeutic Interventions.*

Lifestyle Modifications	Therapeutic Interventions
Weight loss	Oral hypoglycemic agents
Exercise	Metformin
Diet	Thiazolidinediones
Smoking	Alpha-glucosidase inhibitors
Alcohol	Pramlintide
Coffee or tea	GLP-1 agonists
	DDP-4 inhibitors
	Antiobesity agents (eg, orlistat)
	Antihypertensive agents
	Antilipid agents (eg, colesevelam)

Abbreviations: DDP-4, dipeptidyl peptidase 4; GLP-1, glucagon-like peptide 1.

play a significant role in the development of diabetes.[3]

> *The diagnosis of metabolic syndrome appears to play a significant role in the development of diabetes.*

This review focuses on therapeutic options for the prevention of type 2 diabetes in individuals diagnosed with metabolic syndrome. Not all studies reviewed in this section involve a study population with a diagnosis of metabolic syndrome; however, many of these studies evaluate risk factors for metabolic syndrome. In this article, we review the impact of lifestyle and therapeutic interventions (Table 1) on the progression to diabetes in individuals at high risk for metabolic syndrome, such as those with hypertension, dyslipidemia, obesity, and hyperglycemia, particularly impaired fasting glucose (IFG) or impaired glucose tolerance (IGT).

LIFESTYLE INTERVENTIONS

Even in an individual with a genetic predisposition for type 2 diabetes, behavioral factors such as diet and exercise can influence the development of diabetes. Therefore, lifestyle modification such as diet, exercise, and weight loss is the first step in diabetes prevention.

> *Lifestyle modification such as diet, exercise, and weight loss is the first step in diabetes prevention.*

Weight Loss

Several studies have shown that weight loss improves glucose tolerance and prevents progression to diabetes.[4,5] The benefit of lifestyle intervention is highlighted by the Diabetes Prevention Program (DPP).[6] This large multicenter study prospectively followed obese subjects with IGT. Intensive diet and 150 minutes of physical activity per week resulted in an average weight loss of 6.8 kg in the first year, followed by maintenance of this loss for the 3-year duration of the study. A greater reduction in the incidence of diabetes was noted with lifestyle modification compared with placebo (Figure 1), and this finding held true across all participating ethnic groups and for both men and women. In this landmark study, a weight loss of 1 kg resulted in a 16% reduction in the incidence of diabetes.[7] Lifestyle modification worked particularly well for participants aged 60 years and older, reducing their risk of diabetes by 71%. Compared with pharmacological therapy (eg, metformin), lifestyle modification had a greater impact on diabetes prevention and has been associated with improved cardiovascular outcomes.

Diet and Exercise

In addition to weight loss, diet and exercise can independently reduce the risk of diabetes. Several studies have demonstrated a lower risk of diabetes with regular exercise, and this benefit appears to be greater in individuals at high risk for diabetes, such as those with high body mass index

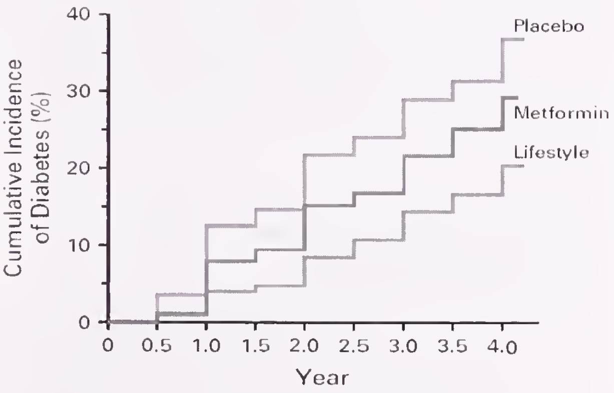


Fig 1. Cumulative incidence of diabetes according to study group. The diagnosis of diabetes was based on the criteria of the American Diabetes Association. The incidence of diabetes differed significantly among the 3 groups ($P < 0.001$ for each comparison). Copyright © 2002 Massachusetts Medical Society. All rights reserved. Reprinted with permission.[6]

(BMI), hypertension, or family history of diabetes.[8,9] Furthermore, greater reductions in diabetes risk are noted with higher intensity and duration of activity. In one study, each 500-kcal increase in the weekly energy expenditure reduced the risk for diabetes by 6%.[10] Even nonintensive leisure activities like shopping, ballroom dancing, and playing volleyball may be beneficial in terms of diabetes prevention.[11]

Dietary composition is another factor that may guide progression to diabetes. Different types of food have different effects on postprandial glycemia. This is known as the glycemic index. The glycemic index and the carbohydrate content of food are incorporated into the determination of glucose load. Glucose load, along with fiber and fat content (eg, saturated, monounsaturated, and polyunsaturated fats), may impact the development of diabetes. In the Nurses' Health Study, a diet high in fiber and polyunsaturated fat and low in glycemic load was associated with a lower risk of developing diabetes.[12] Consumption of food with a high glycemic index increased the risk for diabetes in this study. Other studies have corroborated these findings and found a western diet that is rich in processed meats and fat content increased the risk for diabetes, and this risk was greater in individual with high BMI.[13] However, certain fibers appear to improve insulin sensitivity.[14] In a prospective cohort study of nearly 25,000 men and women, high intake of cereal fiber (eg, whole grains) lowered the risk of diabetes (relative risk [RR], 0.72; 95% confidence interval [CI], 0.56–0.93).[15] This benefit is not seen with soluble fibers found in fruits and vegetables. Consumption of other foods, such as nuts and dairy products, has also been linked to a reduced risk of diabetes, but considering the observational nature of these findings, more studies are needed to clarify their effects on diabetes prevention.[16,17]

Smoking, Alcohol Intake, and Coffee Consumption

Modification of other lifestyle factors such as smoking, alcohol intake, or coffee consumption may also be important in diabetes-prevention efforts. Smoking has been shown to increase blood glucose, impair insulin sensitivity, and increase abdominal fat distribution, resulting in a higher waist-to-hip ratio.[18] The risk of developing diabetes is greater in smokers versus nonsmokers. Even passive or secondhand smoking (RR, 1.35; 95% CI, 1.06–1.71) has been associated with an increased risk of diabetes.[19] This risk increases with quantity and duration of smoking.[20] A recent cohort study found that the HRs of diabetes among former smokers, new quitters, and current smokers compared with never smokers was 1.22 (95% CI, 0.99–1.50), 1.73 (95% CI, 1.19–2.53), and 1.31 (95% CI, 1.04–1.65), respectively.[21] Weight gain associated with smoking cessation may contribute to the development of diabetes, but the risk appears to revert to that of never smokers after over time.[22]

> *Smoking has been shown to increase blood glucose, impair insulin sensitivity, and increase abdominal fat distribution, resulting in a higher waist-to-hip ratio. Even passive or secondhand smoking (RR, 1.35; 95% CI, 1.06–1.71) has been associated with an increased risk of diabetes.*

An inverse diabetes-risk relationship is reported with alcohol consumption. Improved insulin sensitivity and lower diabetes risk have been reported with light to moderate alcohol intake, but not with heavy intake. In the DPP, higher consumption of alcohol was associated with a decrease in insulin resistance and lower rates of diabetes.[23] The RR of diabetes was 0.61 (95% CI, 0.44–0.91) in men who drank 2 to 3 alcoholic beverages per day and 0.4 (95% CI, 0.3–0.6) in women who drink 1 or more alcoholic beverages per day, compared with nondrinkers.[24,25] This benefit has been corroborated by several studies, including a recent meta-analysis.[26]

> *Lower diabetes risk has been reported with light to moderate alcohol intake.*

Consumption of coffee and green tea, independent of caffeine content, may also reduce the risk for type 2 diabetes.[27] Although some studies report

> *Consumption of coffee and green tea, independent of caffeine content, may also reduce the risk for type 2 diabetes.*

an increase in insulin sensitivity with consumption of these products, others report no effect. Various chemicals present in these beverages (eg, magnesium, lignans, chlorogenic acids, and catechins) have been implicated as possible culprits. A recent meta-analysis of 6 large cohort studies found that drinking 3 to 4 cups of coffee per day reduced the risk of diabetes (RR, 0.85; 95% CI, 0.75–0.96) compared with drinking fewer than 2 cups per day. In fact, there was an inverse log-linear relationship between coffee intake and the subsequent risk for diabetes, such that every cup of coffee consumed in 1 day was associated with a 7% decrease in the risk for diabetes.[28]

THERAPEUTIC INTERVENTIONS

When lifestyle interventions fail, pharmacological therapies may be helpful in preventing diabetes.

When lifestyle interventions fail, pharmacological therapies may be helpful in preventing diabetes.

Medications prescribed to treat the various features of the metabolic syndrome (eg, hyperglycemia, obesity, hypertension, and dyslipidemia) may also delay progression to diabetes in high-risk individuals.

Antidiabetic Agents

Several studies document the role of antidiabetic agents such as metformin in diabetes prevention. Metformin has been shown to improve insulin sensitivity and preserve beta-cell function. In the DPP, 2155 individuals with IGT were randomized to receive metformin (850 mg twice a day) or placebo.[6] After an average follow-up of 2.8 years, a 2.5% weight loss and a 31% reduction in the incidence of diabetes was noted in the metformin-treated group compared with placebo. The greatest impact of metformin therapy was primarily observed in individuals under the age of 60 years and those with a BMI >35 kg/m^2.[29] Even after 10 years of continuous metformin use, weight loss was maintained in study participants and the incidence of diabetes was reduced by 18% compared with placebo.[30] In clinical trials, metformin was well tolerated with no serious long-term side effects. Considering the impact of metformin on diabetes prevention, the American Association of Clinical Endocrinologists recently suggested metformin for the treatment of high-risk individuals with IGT or IFG.

Several studies document the role of antidiabetic agents such as metformin in diabetes prevention. In the Diabetes Prevention Program, 2155 individuals with impaired glucose tolerance were randomized to receive metformin (850 mg twice a day) or placebo. After an average follow-up of 2.8 years, a 2.5% weight loss and a 31% reduction in the incidence of diabetes was noted in the metformin-treated group compared with placebo.

Thiazolidinediones have been shown to enhance insulin sensitivity, reduce insulin secretory demand, and preserve beta cell function in diabetes. In the Troglitazone in Prevention of Diabetes (TRIPOD) trial, 266 Hispanic women with previous gestational diabetes were randomized to troglitazone (400 mg/day) or placebo.[31] After median follow-up of 30 months, a 55% reduction in the incidence of diabetes was seen in the treatment arm. In this study, troglitazone improved insulin sensitivity and preserved beta cell function. The metabolic parameters that are altered with therapy persisted long after drug discontinuation. After an 8-month drug-washout period, the annual incidence of diabetes remained significantly lower with troglitazone therapy (3.1% versus 21.2%). Similar findings were noted in the studies involving pioglitazone and

In the Troglitazone in Prevention of Diabetes trial, 266 Hispanic women with previous gestational diabetes were randomized to troglitazone (400 mg/day) or placebo. After median follow-up of 30 months, a 55% reduction in the incidence of diabetes was seen in the troglitazone arm. Troglitazone improved insulin sensitivity and preserved beta cell function.

rosiglitazone, suggesting that the benefits of these drugs may be a class effect.[32] However, safety concerns particularly related to troglitazone (liver failure) and rosiglitazone (heart failure) have limited the use of these agents in diabetes prevention.

Despite efficacy, side effects have also limited the use of alpha-glucosidase inhibitors in diabetes prevention. Acarbose, an alpha-glucosidase inhibitor, reduces the intestinal absorption of carbohydrates and thus postprandial hyperglycemia. Although this reduction is not associated with a delay in progression to frank diabetes in individuals with early diabetes, it does reduce the rate of diabetes development in individuals with IFG or IGT.[33] Once diabetes develops, the reversibility of beta-cell failure with acarbose therapy seems to be limited. In the Study to Prevent Non-Insulin-Dependent Diabetes Mellitus (STOP-NIDDM), 1429 subjects with IGT were randomly assigned to acarbose (100 mg 3 times daily) or placebo.[34] In this intention to treat analysis, the incidence of diabetes was lower with acarbose therapy (32%) compared with placebo (42%) after a 3-year follow-up period. Drug-induced gastrointestinal complaints such as flatulence were responsible for a dropout rate of 25% in the treatment group.

Weight loss associated with some of the newer antidiabetic agents may play a role in diabetes prevention. Pramlintide, an analog of amylin, has been recently discovered and approved for the treatment of insulin-dependent diabetes. Amylin is a peptide released from beta cells in conjunction with insulin after an oral glucose load. Its biological effect includes central inhibition of appetite and glucagon suppression.[35] Multiple studies with diabetic subjects have shown not only improvements in postprandial glycemia with pramlintide, but also weight loss.[36,37] The impact of this agent in nondiabetic subjects is unclear. A recent study of 270 nondiabetics randomized to pramlintide or placebo for 16 weeks found that the treatment group achieved a greater weight loss compared with placebo (6.1 kg versus 2.8 kg, $P < 0.01$).[38] Nausea was the most common adverse effect. Although it is unclear whether weight loss induced by pramlintide translates into diabetes prevention, studies like the DPP have demonstrated a delay in the progression to diabetes with weight reduction. More studies are needed to assess the impact of pramlintide on weight loss and diabetes prevention.

Glucagon-like polypeptide 1 (GLP-1) and gastric inhibitory polypeptide (GIP) are produced by the distal intestine in response to carbohydrate-rich meals. They have a very short half-life and are rapidly degraded by dipeptidyl peptidase 4 (DPP-4). They inhibit glucagon secretion and delay gastric emptying.[39] In vivo and in vitro studies indicate that these peptides protect against beta-cell apoptosis and increase beta-cell mass.[40] Currently, several drugs that impact this pathway have been approved for the treatment of diabetes. Both GLP-1 agonists (eg, exenatide and liraglutide) and DPP-4 inhibitors (eg, sitagliptin and saxagliptin) have been shown to improve glycemic control in diabetics.[41,42] In addition, GLP-1 agonists may induce weight loss.[42] Limited data are available on the effect of these agents in the treatment of metabolic syndrome or diabetes prevention.

In studies involving insulin-resistant rats, exenatide has been shown to increase glucose-stimulated insulin secretion and to increase beta-cell mass.[43] Even after treatment discontinuation, improved glucose tolerance and insulin sensitivity were reported. Weight loss associated with exenatide treatment in diabetic subjects has prompted off-label use of this agent in high-risk individuals. In a retrospective chart review of 299 obese patients treated with exenatide, a weight loss of 6% was noted after 16 weeks.[44] In clinical trials, another GLP-1 agonist, liraglutide, induced greater weight loss and improved glucose tolerance after 20 weeks of treatment versus placebo.[45] In fact, up to a 96% reduction in the prevalence of IGT was noted with liraglutide treatment. Although neither study assessed glucose or insulin levels, the weight-loss and beta-cell-protecting effect of GLP-1 agonists may be a path to achieve insulin sensitivity in patients treated with these agents. More studies are needed to elucidate the role of these agents, as well as DDP-4 inhibitors, on diabetes prevention.

> *The weight-loss and beta-cell-protecting effect of GLP-1 agonists may be a path to achieve insulin sensitivity in patients.*

Treatment for Obesity

As our understanding of the role of weight loss in diabetes expands, pharmacological agents aimed at weight loss may guide the future of diabetes prevention. Approved antiobesity agents work by altering fat digestion or by causing early satiety. Orlistat, an antiobesity medication, is a potent inhibitor of intestinal lipase, reducing absorption of dietary fat (TG).[46] In a randomized trial, orlistat caused an 11% weight loss in the first year of therapy.[47] Although patients regained weight in subsequent years, orlistat treatment maintained a 6.9% loss at the end of the 4-year study period, compared with a 4.1%

Pharmacological agents aimed at weight loss may guide the future of diabetes prevention.

loss maintained by those on the placebo arm of the study. In obese subjects with IGT, a 37% reduction in the incidence of diabetes is noted with treatment. Despite these benefits, high dropout rates due to gastrointestinal side effects (ie, cramps, flatus) were reported in this study and have limited the use of this agent in the clinical setting. However, sympathomimetic agents that suppress appetite are better tolerated. Some of these agents, such as phentermine and diethylpropion, have a potential for abuse, and therefore are not considered appropriate for long-term use. Sibutramine, an agent that decreases reuptake of serotonin and noradrenalin, has been approved by the Food and Drug Administration for weight management.[48] Weight loss of 9.95 kg after 1 year of treatment was reported in a recent meta-analysis.[49] The effect of sibutramine on weight loss appears to be greater when combined with lifestyle intervention. In addition, sibutramine may reduce insulin resistance independent of weight loss in obese subjects and improve glycemic control in diabetics.[50] Changes in lipid parameters (ie, lower TG and low-density lipoprotein cholesterol [LDL-C] and higher HDL-C) and elevations of blood pressure were also reported with treatment. These changes appear to be dose dependent; therefore, careful assessment and monitoring of blood pressure is required prior to initiation and with continuation of therapy.

Furthermore, weight loss has also been associated with drugs approved for the treatment of other medical conditions, such as depression (eg, fluoxetine, bupropion) and seizure disorders (eg, topiramate, zonisamide). Although these agents are not approved for use in weight management, weight loss has been reported in clinical trials utilizing these agents, and more studies are needed to clarify the impact of these agents on glycemic control. Several studies have evaluated the effect of combining different classes of agents to maximize weight loss and improve glycemic control. However, the results have not been fruitful. Combination orlistat and sibutramine did not produce more weight loss than sibutramine alone.[52,53] Qnexa, a medication that combines phentermine and topiramate, was recently pull from the market due to safety concerns despite an average weight loss of 9.2% and improved glycemic control after 6 months.[51] Psychiatric and cognitive side effects, potential risk for birth defects, and inadequate data on cardiovascular safety were some of the concerns sited by the FDA.

Bariatric surgery is playing a greater role in weight management and diabetes prevention.

With limited pharmacological options, bariatric surgery is playing a greater role in weight management and diabetes prevention in morbidly obese individuals with comorbidities such as hypertension, IGT, and dyslipidemia. Several studies document weight loss associated diabetes remission with bariatric surgery.[54,55] However, immediate changes in levels of intestinal peptides may also improve insulin resistance and glycemic control, and these changes appear to occur independent of weight loss.[56,57]

Antilipid and Antihypertensive Drugs

Antilipid and antihypertensive agents are not routinely recommended for the prevention of diabetes. However, high-risk individuals, such as those with metabolic syndrome, may receive these medications to control their blood pressure and treat their lipid disorders. Preliminary analysis of studies involving these agents suggests that some of the medications prescribed to treat these conditions may also have a role in diabetes prevention.

Inhibition of angiotensin II by angiotensin-converting enzyme inhibitors or angiotensin receptor blockers may reduce insulin resistance.

Angiotensin II has an inhibitory effect on insulin signaling, production of vascular nitric oxide, and vasodilatation.[58] Thus, inhibition of angiotensin II by angiotensin-converting enzyme inhibitors or angiotensin receptor blockers may reduce insulin resistance. A meta-analysis of trials involving angiotensin-converting enzyme inhibitors and angiotensin receptor blockers found a 25% decrease in the incidence of diabetes with these agents.[59] These results are supported by a recent publication involving more than 9000 subjects with IGT randomized to valsartan or placebo for 5 years.[60] Reduction in the development of diabetes was reported in the valsartan arm (RR, 0.86; 95% CI, 0.80–0.92, $P < 0.01$). However, a prior randomized control trial involving ramipril, the Diabetes Reduction Assessment with Ramipril and Rosiglitazone Medication (DREAM) trial, did not find a decrease in the incidence of new diabetes over 3 years of follow-up (RR, 0.91; 95% CI, 0.81–1.03).[61] The discrepancy between these studies is not entirely clear. Regardless, studies supporting improved insulin sensitivity with these agents are interesting and the issue needs further clarification.[62]

> *Lipid-lowering medications play an important role in the management of metabolic syndrome.*

Lipid-lowering medications play an important role in the management of metabolic syndrome. Statins are routinely prescribed for cardiovascular prevention. In the West of Scotland Coronary Prevention Study, pravastatin decreased the incidence of diabetes by increasing insulin sensitivity, but this association was not supported by other studies.[63] In fact, other statins like rosuvastatin were found to worsen insulin sensitivity.[64] Meta-analysis of studies involving statins found a small increase in risk of diabetes.[65] In contrast, the bile acid sequestrant colesevelam hydrochloride modestly improves blood glucose levels and therefore has been approved by the Food and Drug Administration for diabetes management.[66] Although the exact mechanism is unknown, one hypothesis focuses on the effect of this agent on glucose absorption. A post-hoc analysis of laboratory safety data found treatment with colesevelam reduced fasting plasma glucose in a subgroup of 88 subjects with IFG compared with placebo at 24 weeks.[67] Even though this is a small study population, it suggests that colesevelam may be effective in preventing diabetes.

Fibrates are nuclear proliferator-activated receptor alpha agonists critical in lowering TG. In diet-induced obese mice, fenofibrate improved glucose tolerance by decreasing visceral adipocyte size and quantity.[68] However, clinical trials involving fibrates have not reported weight reduction or improved insulin sensitivity.[69,70] However, one post-hoc analysis found a 30% reduction in the incidence of diabetes with bezafibrate treatment.[71] More studies are needed to further clarify the role of these agents.

CONCLUSION

As our understanding of metabolic syndrome continues to progress, preventing complications like diabetes is becoming the focus of management strategies. In adults with metabolic syndrome, data support the fact that both lifestyle and therapeutic interventions can influence progression to diabetes. Diet and exercise programs resulting in weight loss should be considered the first line in diabetes prevention. Pharmacological intervention with metformin seems promising. However, the long-term effectiveness of other therapeutic interventions is less clear. More studies are needed to determine the cost-effectiveness of diabetes-prevention strategies and to help guide prevention efforts in metabolic syndrome.

DISCLOSURES

Potential conflict of interest: Gopalakrishnan: Roche and Amgen (PI in clinical trial program). Lukacova-Zib: Roche and Amgen (Co-Investigator in clinical trial program).

REFERENCES

1. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998; 15: 539–553.
2. Alberti KG, Eckel RH, Grundy SM, et al. International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; International Association for the Study of Obesity. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009; 120: 1640–1645.
3. Klein BE, Klein R, Lee KE. Components of the metabolic syndrome and risk of cardiovascular disease and diabetes in Beaver Dam. *Diabetes Care* 2002; 25: 1790–1794.
4. Long SD, O'Brien K, MacDonald KG Jr, et al. Weight loss in severely obese subjects prevents the progression of impaired glucose tolerance to type I diabetes: a longitudinal interventional study. *Diabetes Care* 1994; 17: 372–375.
5. Wing RR, Venditti E, Jakicic JM, et al. Lifestyle intervention in overweight individuals with family history of diabetes. *Diabetes Care* 1998; 21: 350–359.
6. Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002; 346: 393–403.
7. Hamman RF, Wing RR, Edelstein SL, et al. Effect of weight loss with lifestyle intervention on risk of diabetes. *Diabetes Care* 2006; 29: 2102–2107.
8. Lynch J, Helmrich SP, Lakka TA, et al. Moderately intense physical activities and high levels of

cardiorespiratory fitness reduce the risk of non-insulin-dependent diabetes mellitus in middle-aged men. *Arch Intern Med* 1996; 156: 1307–1314.

9. Hu FB, Sigal RJ, Rich-Edwards JW, et al. Walking compared with vigorous physical activity and risk of type 2 diabetes in women: a prospective study. *JAMA* 1999; 282: 1433–1439.
10. Helmrich SP, Ragland DR, Leung RW, et al. Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. *N Engl J Med* 1991; 325: 147–152.
11. Pan XR, Li GW, Hu YH, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the Da Qing IGT and Diabetes Study. *Diabetes Care* 1997; 20: 537–544.
12. Shai I, Jiang R, Manson JE, et al. Ethnicity, obesity, and risk of type 2 diabetes in women: a 20-year follow-up study. *Diabetes Care* 2006; 29: 1585–1590.
13. Van Dam RM, Rimm EB, Willett WC, et al. Dietary patterns and risk for type 2 diabetes mellitus in US men. *Ann Intern Med* 2002; 136: 201–209.
14. Weickert MO, Mohlig M, Schofl C, et al. Cereal fiber improves whole-body insulin sensitivity in overweight and obese women. *Diabetes Care* 2006; 29: 775–780.
15. Schulze MB, Schulz M, Heidemann C, et al. Fiber and magnesium intake and incidence of type 2 diabetes: a prospective study and meta-analysis. *Arch Intern Med* 2007; 167: 956–965.
16. Pereira MA, Jacobs DR Jr, Van Horn L, et al. Dairy consumption, obesity, and the insulin resistance syndrome in young adults: the CARDIA Study. *JAMA* 2002; 287: 2081–2089.
17. Jaing R, Manson JE, Stampfer MJ, et al. Nut and peanut butter consumption and risk of type 2 diabetes in women. *JAMA* 2002; 288: 2554–2560.
18. Frati AC, Iniestra F, Ariza CR. Acute effect of cigarette smoking on glucose tolerance and other cardiovascular risk factors. *Diabetes Care* 1996; 19: 112–118.
19. Houston TK, Person SD, Pletcher MJ, et al. Active and passive smoking and development of glucose intolerance among young adults in prospective cohort: CARDIA study [published correction appears in *BMJ* 2006; 332: 1298]. *BMJ* 2006; 332: 1064–1069.
20. Manson JE, Ajani UA, Liu S, et al. A prospective study of cigarette smoking and incidence of diabetes mellitus among US male physicians. *Am J Med* 2000; 109: 538–542.
21. Yeh HC, Duncan BB, Schmidt MI, et al. Smoking, smoking cessation, and risk for type 2 diabetes mellitus: a cohort study. *Ann Intern Med* 2010; 152: 10–17.
22. Wannamethee SG, Shaper AG, Perry IJ; British Regional Heart Study. Smoking as a modifiable risk factor for type 2 diabetes in middle-aged men. *Diabetes Care* 2001; 24: 1590–1595.
23. Crandall JP, Polsky S, Howard AA, et al. Alcohol consumption and diabetes risk in the Diabetes Prevention Program. *Am J Clin Nutr* 2009; 90: 595–601.
24. Rimm EB, Chan J, Stampfer MJ, et al. Prospective study of cigarette smoking, alcohol use, and the risk of diabetes in men. *BMJ* 1995; 310: 555–559.
25. Stampfer MJ, Colditz GA, Willett WC, et al. A prospective study of moderate alcohol drinking and risk of diabetes in women. *Am J Epidemiol* 1988; 128: 549–558.
26. Koppes LL, Dekker JM, Hendriks HF, et al. Moderate alcohol consumption lowers the risk of type 2 diabetes: a meta-analysis of prospective observational studies. *Diabetes Care* 2005; 28: 719–725.
27. Huxley R, Lee CM, Barzi F, et al. Coffee, decaffeinated coffee, and tea consumption in relation to incident type 2 diabetes mellitus: a systematic review with meta-analysis. *Arch Intern Med* 2009; 169: 2053–2063.
28. Van Dam RM, Willett WC, Manson JE, Hu FB. Coffee, caffeine and risk of type 2 diabetes: a prospective cohort study in younger and middle-aged US women. *Diabetes Care* 2006; 29: 398–403.
29. Herman WH, Hoerger TJ, Brandle M, et al. The cost-effectiveness of lifestyle modification or metformin in preventing type 2 diabetes in adults with impaired glucose tolerance. *Ann Intern Med* 2005; 142: 323–332.
30. Diabetes Prevention Program Research Group; Knowler WC, Fowler SC, Hamman RF, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study [published correction appears in *Lancet* 2009; 374: 2054]. *Lancet* 2009; 374: 1677–1686.
31. Buchanan TA, Xiang AH, Peters RK, et al. Preservation of pancreatic beta-cell function and prevention of type 2 diabetes by pharmacological treatment of insulin resistance in high-risk Hispanic women. *Diabetes* 2002; 51: 2796–2803.
32. Xiang AH, Peters RK, Kjos SL, et al. Effect of pioglitazone on pancreatic beta-cell function and diabetes risk in Hispanic women with prior gestational diabetes. *Diabetes* 2006; 55: 517–522.
33. Kirkman MS, Shankar RR, Shankar S, et al. Treating postprandial hyperglycemia does not appear to delay progression of early type 2 diabetes: the Early Diabetes Intervention Program. *Diabetes Care* 2006; 29: 2095–2101.
34. Chiasson JL, Josse RG, Gomis R, et al. Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomised trial. *Lancet* 2002; 359: 2072–2707.
35. Roth JD, Hughes H, Kendall E, et al. Antiobesity effects of the beta-cell hormone amylin in diet-induced obese rats: effects on food intake, body weight, composition, energy expenditure, and gene expression. *Endocrinology* 2006; 147: 5855–5864.
36. Brown A, Desai M, Taneja D, et al. Managing highly insulin-resistant diabetes mellitus: weight loss approaches and medical management. *Postgrad Med* 2010; 122: 163–171.
37. Ryan G, Briscoe TA, Jobe L. Review of pramlintide as adjunctive therapy in treatment of type 1 and type 2 diabetes. *Drug Des Devel Ther* 2009; 2: 203–214.
38. Smith SR, Aronne LJ, Burns CM, et al. Sustained weight loss following 12-month pramlintide treatment as an adjunct to lifestyle intervention in obesity. *Diabetes Care* 2008; 31: 1816–1823.
39. Baggio LL, Drucker DJ. Biology of incretins: GLP-1 and GIP. *Gastroenterology* 2007; 132: 2131–2157.
40. Baggio LL, Drucker DJ. Therapeutic approaches to preserve islet mass in type 2 diabetes. *Annu Rev Med* 2006; 57: 265–281.
41. Hermansen K, Kipnes M, Luo E, et al. Sitagliptin Study Group. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, in patients with type 2 diabetes mellitus inadequately controlled on glimepiride alone or on glimepiride and metformin. *Diabetes Obes Metab* 2007; 9: 733–745.

42. Buse JB, Klonoff DC, Nielsen LL, et al. Metabolic effects of two years of exenatide treatment on diabetes, obesity, and hepatic biomarkers in patients with type 2 diabetes: an interim analysis of data from the open-label, uncontrolled extension of three double-blind, placebo-controlled trials. *Clin Ther* 2007; 29: 139–153.
43. Xu G, Stoffers DA, Habener JF, Bonner-Weir S. Exendin-4 stimulates both beta-cell replication and neogenesis, resulting in increased beta-cell mass and improved glucose tolerance in diabetic rats. *Diabetes* 1999; 48: 2270–2276.
44. Bhushan R, Elkind-Hirsch KE, Bhushan M, et al. Exenatide use in the management of metabolic syndrome: a retrospective database study. *Endocr Pract* 2008; 14: 993–999.
45. Astrup A, Rossner S, Van Gaal L, et al. Effects of liraglutide in the treatment of obesity: a randomised, double-blind, placebo-controlled study [published correction appears in *Lancet* 2010; 375: 984]. *Lancet* 2009: 7; 374: 1606–1616.
46. Kiortsis DN, Filippatos TD, Elisaf MS. The effects of orlistat on metabolic parameters and other cardiovascular risk factors. *Diabetes Metab* 2005; 31: 15–22.
47. Torgerson JS, Hauptman J, Boldrin MN, Sjöström L. XENical in the prevention of diabetes in obese subjects (XENDOS) study: a randomized study of orlistat as an adjunct to lifestyle changes for the prevention of type 2 diabetes in obese patients [published correction appears in *Diabetes Care* 2004; 27: 856]. *Diabetes Care* 2004; 27: 155–161.
48. Astrup A, Hansen DL, Lundsgaard C, Toubro S. Sibutramine and energy balance. *Int J Obes Relat Metab Disord* 1998; 22(suppl 1): S30–S35.
49. Li Z, Maglione M, Tu W, et al. Meta-analysis: pharmacologic treatment of obesity. *Ann Intern Med* 2005; 142: 532–546.
50. Kim DM, Yoon SJ, Ahn CW, et al. Sibutramine improves fat distribution and insulin resistance, and increases serum adiponectin levels in Korean obese nondiabetic premenopausal women. *Diabetes Res Clin Pract* 2004; 66(suppl 1): S139–S144.
51. 17th Annual European Congress on Obesity (ECO), Amsterdam, The Netherlands. Results from Phase 3 EQUATE Trial of VIVUS' Qnexa. http://ir.vivus.com/releasedetail.cfm?ReleaseID=382339. Published May 7, 2009.
52. Wadden TA, Berkowitz RI, Womble LG, et al. Effects of sibutramine plus orlistat in obese women following 1 year of treatment by sibutramine alone: a placebo-controlled trial. *Obes Res* 2000; 8: 431–437.
53. Kaya A, Aydin N, Topsever P, et al. Efficacy of sibutramine, orlistat and combination therapy on short-term weight management in obese patients. *Biomed Pharmacother* 2004; 58: 582–587.
54. Sjöström L, Lindroos AK, Peltonen M, et al. Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. *N Engl J Med* 2004; 351: 2683–2693.
55. Dixon JB, O'Brien PE, Playfair J, et al. Adjustable gastric banding and conventional therapy for type 2 diabetes: a randomized controlled trial. *JAMA* 2008; 299: 316–323.
56. Rizzello M, Abbatini F, Casella G, et al. Early postoperative insulin-resistance changes after sleeve gastrectomy. *Obesity Surg* 2010; 20: 50–55.
57. Rubino F, Gagner M, Gentileschi P, et al. The early effect of the Roux-en-Y gastric bypass on hormones involved in body weight regulation and glucose metabolism. *Ann Surg* 2004; 240: 236–242.
58. Folli F, Kahn CR, Hansen H, et al. Angiotensin II inhibits insulin signaling in aortic smooth muscle cells at multiple levels: a potential role for serine phosphorylation in insulin/angiotensin II crosstalk. *J Clin Invest* 1997; 100: 2158–2169.
59. Abuissa H, Jones PG, Marso SP, O'Keefe JH Jr. Angiotensin-converting enzyme inhibitors or angiotensin receptor blockers for prevention of type 2 diabetes: a meta-analysis of randomized clinical trials. *J Am Coll Cardiol* 2005; 46: 821–826.
60. McMurray JJ, Holman RR, Haffner SM, et al. NAVIGATOR Study Group. Effect of valsartan on the incidence of diabetes and cardiovascular events [published correction appears in *N Engl J Med* 2010; 362: 1748]. *N Engl J Med* 2010; 362: 1477–1490.
61. Bosch J, Yusuf S, Gerstein HC, et al. DREAM Trial Investigators. Effect of ramipril on the incidence of diabetes. *N Engl J Med* 2006; 355: 1551–1562.
62. Taniyama Y, Hitomi H, Shah A, et al. Mechanisms of reactive oxygen species-dependent downregulation of insulin receptor substrate-1 by angiotensin II. *Arterioscler Thromb Vasc Biol* 2005; 25: 1142–1147.
63. Freeman DJ, Norrie J, Sattar N, et al. Pravastatin and the development of diabetes mellitus: evidence for a protective treatment effect in the West of Scotland Coronary Prevention Study. *Circulation* 2001; 103: 357–362.
64. Ridker PM, Danielson E, Fonseca FA, et al. JUPITER Study Group. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N Engl J Med* 2008; 359: 2195–2207.
65. Baker WL, Talati R, White CM, Coleman CI. Differing effect of statins on insulin sensitivity in non-diabetics: a systematic review and meta-analysis. *Diabetes Res Clin Pract* 2010; 87: 98–107.
66. Fonseca VA, Rosenstock J, Wang AC, et al. Colesevelam HCl improves glycemic control and reduces LDL cholesterol in patients with inadequately controlled type 2 diabetes on sulfonylurea-based therapy. *Diabetes Care* 2008; 31: 1479–1484.
67. Handelsman Y, Abby SL, Jin X, et al. Colesevelam HCl improves fasting plasma glucose and lipid levels in patients with prediabetes. *Postgrad Med* 2009; 121: 62–69.
68. Jeong S, Yoon M. Fenofibrate inhibits adipocyte hypertrophy and insulin resistance by activating adipose PPARalpha in high fat diet-induced obese mice. *Exp Mol Med* 2009; 41: 397–405.
69. Okopień B, Krysiak R, Herman ZS. Effect of short-term fenofibrate treatment on circulating markers of inflammation and hemostasis in patients with impaired glucose tolerance. *J Clin Endocrinol Metab* 2006; 91: 1770–1778.
70. Belfort R, Berria R, Cornell J, Cusi K. Fenofibrate reduces systemic inflammation markers independent of its effect on lipid and glucose metabolism in patients with the metabolic syndrome. *J Clin Endocrinol Metab* 2010; 95: 829–836.
71. Tenenbaum A, Motro M, Fisman EZ, et al. Peroxisome proliferator-activated receptor ligand bezafibrate for prevention of type 2 diabetes mellitus in patients with coronary artery disease. *Circulation* 2004; 109: 2197–2202.

Mount Sinai Journal of Medicine 77:533–548, 2010

Weight Management in Type 2 Diabetes Mellitus

Amulya T. Siram, MD, Robert Yanagisawa, MD, and Maria Skamagas, MD

Department of Medicine, Division of Endocrinology, Diabetes, and Bone Diseases, Mount Sinai School of Medicine, New York, NY

OUTLINE

Obesity: Independent Risk Factor for Complications and Mortality in Type 2 Diabetes Mellitus
Lifestyle Modification: Weight Management and Glycemic Control
Diabetes Medications: Glycemic Control, Weight Gain, and Hypoglycemia
Pharmacologic Management of Type 2 Diabetes: Effect on Glycemic Control and Weight Management
Diabetes Mellitus Management Strategies
Conclusion

ABSTRACT

Obesity is a well known risk factor for type 2 diabetes mellitus. Individuals with type 2 diabetes mellitus are at risk for weight gain as a result of multiple influences, including sedentary lifestyle, high-calorie diet, diabetes medications, sociocultural factors, chronic medical and psychiatric illnesses, and a dysregulated enteroendocrine axis. Because both diabetes mellitus and obesity predispose patients to abnormal cardiometabolic profiles and increased cardiovascular disease, management of diabetes mellitus should focus on weight management and optimizing cardiometabolic parameters, concomitant with glycemic control. Lifestyle modification incorporating healthy, calorie-appropriate diets and increased physical activity, in addition to metformin, are central components to diabetes management and weight management. These interventions have been shown to improve body weight, glycemic control, and overall cardiometabolic profile. The weight-neutral and weight-losing diabetes medications include metformin, alpha-glucosidase inhibitors, glucagon-like peptide-1 analogs, dipeptidyl peptidase-4 inhibitors, and amylin analogs. It is essential that providers understand the metabolic and weight effects of diabetes medications in order to develop strategies for managing diabetes mellitus while helping patients maintain or lose weight in order to improve their overall health outcomes. ***Mt Sinai J Med 77:533–548, 2010.*** **©** *2010 Mount Sinai School of Medicine*

Key Words: enteroendocrine axis, metabolic syndrome, obesity, type 2 diabetes mellitus, weight management.

Address Correspondence to:

Maria Skamagas, MD
Department of Medicine
Division of Endocrinology,
Diabetes, and Bone Diseases
Mount Sinai School of Medicine
New York, NY
Email: maria.skamagas@mssm.edu

Obesity is a well-known risk factor for type 2 diabetes mellitus (T2DM).[1–3] Conversely, individuals with T2DM are at risk for weight gain as a result of multiple influences, including sedentary lifestyle, high-calorie diets, diabetes medications, sociocultural factors, chronic medical and psychiatric illnesses, and a dysregulated enteroendocrine axis. Obesity predisposes individuals to abnormal cardiometabolic profiles: increased waist circumference with increased visceral adiposity and inflammatory adipokines, insulin resistance with impaired glucose tolerance or diabetes mellitus (DM), elevated triglycerides (TG), decreased high-density lipoprotein (HDL) cholesterol, and hypertension.[4–6] Ultimately, DM and obesity contribute to increased cardiovascular disease (CVD), microvascular disease, and death. Although it is clear that achieving glycemic control reduces microvascular disease in the eyes, kidneys, and peripheral nerves, it has not been definitively demonstrated that glycemic control improves macrovascular outcomes, including myocardial infarction, stroke, and peripheral artery disease. Weight loss has been shown time and time again to yield substantial improvements in blood glucose, lipid levels, and blood pressure (BP), each

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20208

© 2010 Mount Sinai School of Medicine

of which is a major risk factor for CVD.[7] Thus, weight loss and weight maintenance with amelioration of the metabolic milieu must attain a similar priority as glycemic control in order to achieve improved cardiovascular (CV) and microvascular health.

Paradoxically, many of the pharmacologic agents used to optimize glycemic control and minimize the microvascular complications of DM often bring about weight gain and worsening of preexisting obesity. Insulin, sulfonylureas, thiazolidinediones (TZDs), and meglitinides are potent means of controlling glucose, but simultaneously cause significant weight gain. Management of DM includes continuous and repeated emphasis on nutrition and physical activity, a multipronged team involving nutritionists and diabetes educators, and a pharmacologic strategy that minimizes weight gain. The weight-neutral and weight-losing diabetes medications include metformin, alpha-glucosidase inhibitors (AGIs), glucagon-like peptide-1 (GLP-1) analogs, dipeptidyl peptidase-4 (DPP-4) enzyme inhibitors, and amylin analogs.

Insulin, sulfonylureas, thiazolidinediones, and meglitinides are potent means of controlling glucose, but simultaneously cause significant weight gain. The weight-neutral and weight-losing diabetes medications include metformin, alpha-glucosidase inhibitors, glucagon-like peptide-1 analogs, dipeptidyl peptidase-4 enzyme inhibitors, and amylin analogs.

OBESITY:
INDEPENDENT RISK FACTOR FOR COMPLICATIONS AND MORTALITY IN TYPE 2 DIABETES MELLITUS

Several studies illustrate the importance of obesity as an independent risk factor for CVD and mortality in T2DM. In the Framingham Heart Study, subjects with DM from the original and offspring cohorts were analyzed for the cumulative risk of CVD beginning at age 50.[8] Normal-weight women and men with DM had a 30-year lifetime risk of CVD of 55% and 79% respectively, compared with 79% and 87% in obese women and men with DM. In another observational study, conducted by the Swedish National Diabetes Registrar, 13,087 patients with T2DM and no known CVD were followed for 6 years. Being overweight and obese increased the risk of both CVD (overweight: odds ratio [OR], 1.24; obese: OR, 1.44) and mortality (overweight: OR, 1.16; obese: OR, 1.7) independent of other CV risk factors.[9]

Several studies illustrate the importance of obesity as an independent risk factor for cardiovascular disease and mortality in type 2 diabetes mellitus.

Metabolic syndrome is an independent risk factor for DM complications. In the Verona Diabetes Complications Study, patients with T2DM were assessed for the presence of metabolic syndrome and development of CV morbidity and mortality over 4.5 years. At the start of the study, 32% had CVD, which was more prevalent among subjects who also had the metabolic syndrome (32.9% versus 17.8%, $P = 0.0005$). Among subjects without CVD at the start of the study, incident CV events were 5-fold greater in those with metabolic syndrome as compared with those patients without it (19.9% versus 3.9%, $P < 0.001$).[10] In a cross-sectional study of 8497 patients with DM randomly chosen from outpatient clinics in Italy, the presence of metabolic syndrome was an independent statistical indicator of the presence of both microvascular and CV complications.[11] These studies illustrate the contribution of excess weight and metabolic syndrome to end-organ disease in T2DM. Therefore, weight management should be an integral component of comprehensive DM management.

LIFESTYLE MODIFICATION: WEIGHT MANAGEMENT AND GLYCEMIC CONTROL

Lifestyle interventions including nutrition, exercise, and behavioral modification to induce weight loss may significantly improve an individual's cardiometabolic profile and reduce the risk of microvascular disease, CVD, and mortality. Even modest weight loss of 5% to 10% of total body weight (10–20 pounds in a 200-pound individual) improves glucose levels, lipid profiles, and BP in individuals with prediabetes and DM.[7,12–16]

Even modest weight loss of 5% to 10% of total body weight (10–20 pounds in a 200-pound individual) improves glucose levels, lipid profiles, and blood pressure in individuals with prediabetes and diabetes.

A meta-analysis of weight-loss intervention studies involving a total of 4659 participants undergoing diet, exercise, and/or behavioral interventions found that any given intervention produced a 3.1% weight loss from baseline (equivalent to 1.7 kg, $P < 0.05$) over 1 to 2 years of follow-up as compared with usual care.[17]

Poor nutritional habits and excess consumption of calories are major components of weight gain, worsening cardiometabolic profiles and glycemic control in patients with T2DM. Individuals with DM may not be aware of the excess calories they consume[18] and may not adhere to prescribed diets.[19] The effect of diet on weight and glycemic control is most clearly seen when patients are placed on a restricted-calorie diet. A meta-analysis of 18 studies in obese T2DM subjects found that individuals who lost weight on low-calorie diets had reductions in glucose levels. Ten-percent weight loss from baseline–ie, 10 kg weight loss over 12 weeks, equivalent to 2 pounds per week–lowered glucose values by approximately 25%. Greater reductions in weight correlated with greater reductions in fasting plasma glucose. Furthermore, weight loss reduced BP, both systolic and diastolic, by 8%; serum cholesterol by 9%; and serum TG by 27%.[16] A similar improvement in glycemic control and reduction in insulin requirements is often encountered in hospitalized patients with T2DM who eat a calorie-controlled diet as compared with an unrestricted diet at home. Thus, excess calorie intake plays a major role in glucose dysregulation and is an important target for management of body weight and DM.

Along with poor dietary compliance, studies have shown that many patients with T2DM lead sedentary lifestyles with minimal physical activity.[20,21] Exercise is an important part of DM management, as it may increase insulin sensitivity, improve body weight, and improve cardiometabolic parameters. A study of inactive subjects with T2DM, randomized to aerobic training, resistance training, both forms of training, or to a control group for 22 weeks, showed reduction in body weight and body mass index (BMI) in the aerobic group compared with controls. There was also a reduction in waist circumference in both training groups as compared with controls, but no difference in visceral adiposity. Glycated hemoglobin A_{1c} (HbA_{1c}; A_{1c}) reduction was about 0.5% for each training group compared with controls; the combined aerobic and resistance training achieved a greater A_{1c} reduction than either exercise alone.[22] A meta-analysis of 14 randomized controlled exercise studies (12 aerobic training studies, 2 resistance training studies) in T2DM demonstrated that exercise reduced A_{1c} by 0.66% ($P < 0.001$), but there was no significant change in BMI between exercise and control groups.[23] Therefore, increased physical activity, with or without a resultant change in body weight, improves glycemic control in T2DM.

Adherence to formal exercise programs tends to be poor, but even home-based moderate-intensity exercise on a routine basis has been shown to improve body weight and cardiometabolic profile. For instance, home-based resistance training was as effective as gymnasium-based training in weight loss and glucose improvements.[24] In addition, 60 minutes of brisk walking 3 times per week compared with a formal medical exercise program was equivalent in improving glycemic control and CV risk factors.[25]

Maintaining weight loss may be more challenging in T2DM. In one study, similar weight loss was achieved by obese women with and without DM enrolled in a behavioral weight-loss program (7.4 kg versus 6.4 kg), but by 1 year follow-up, the women with DM gained back more weight than their nondiabetic counterparts (5.4 kg versus 1.0 kg).[26] Interestingly, 24-hour energy expenditure measured in whole-body respiratory chambers is the same in overweight or obese T2DM subjects compared with overweight or obese controls.[27] When adjusted for various parameters including body composition, subjects with DM actually had an increased basal metabolic rate.

Intensive lifestyle modification produces significant weight loss and improved CV risk factors as evidenced in the Look AHEAD (Action for Health in Diabetes) study, the largest randomized study of lifestyle intervention in overweight T2DM individuals. The study randomized 5145 subjects to 1 year of an intensive lifestyle intervention versus a comparator group of DM support and education.[28] The intensive group was assigned to 42 sessions (75% group and 25% individual sessions) over 1 year in which they received diet and physical activity counseling, with a focus on calorie restriction: portion control, liquid meal replacements and frozen food entrées, or structured meal plans. Activity included home-based exercise with a goal target of 175 minutes per week, which included walking and other moderate-intensity exercise. The comparator group received 3 group educational sessions on diet, exercise, and social support. The intensive subjects each attended a mean of 35 sessions, exercised a mean of 137 minutes per week, and consumed 361 meal replacements. After 1 year of participation, the intensive group achieved 8.6% weight loss from baseline versus 0.7% in the comparator group ($P < 0.001$) and reduction in A_{1c} of 0.64% versus 0.14% ($P < 0.001$). The intensive group also demonstrated significant improvements in BP, TG, urine albumin to creatinine ratio, and prevalence of metabolic syndrome. They required fewer medications to control cardiac risk factors and had lower overall cost of medications.[29]

Lifestyle modification and weight loss is an achievable goal in people with T2DM, but maintaining weight loss requires continued efforts.

> *Lifestyle modification and weight loss is an achievable goal in people with type 2 diabetes mellitus, but maintaining weight loss requires continued efforts.*

A multipronged team approach involving physicians, registered dietitians, certified diabetes educators, and physician extenders is necessary to meet the patient's multiple needs. Educational sessions that support healthy eating and physical activity improve an individual's ability to lose weight.[16] A meta-analysis showed that self-management education improved A_{1c} compared with controls, and that a decrease in A_{1c} of 1% was seen for each additional 23.6 hours of contact that a patient had with an educator.[30] However, this benefit declined 1–3 months after the educational interventions ceased, suggesting that constant reinforcement and counseling is essential in patient management.

When addressing lifestyle modification, practitioners should be aware of the barriers to changing nutritional and physical activity habits in individuals with DM. These include medical, cultural, psychological, social, and financial factors that may drive patient behavior. Some examples include limitations on physical activity due to physical deconditioning, heart disease, and arthritis; cultural eating patterns that favor carbohydrate-rich and energy-dense foods; fear of hypoglycemia, which causes reactive eating and diminished activity; mental illness, including depression and schizophrenia; inability to afford costly DM medications; and lack of social support in the home. To assist the patient in making lifestyle changes, practitioners must openly discuss the difficulties that the patient is experiencing.

Furthermore, patients should be counseled on realistic and concrete weight-loss goals and lifestyle changes. Although there is no clear evidence, a generally safe and realistic weight-loss rate is 1–2 pounds per week until the goal weight is reached. A net negative energy balance of approximately 3500 kcal is required to lose 1 pound of body fat.[31] Therefore, a net negative energy balance of 500 kcal per day for 1 week results in a 1-pound weight loss. This may be achieved through reduced calorie intake and increased physical activity. Other simple lifestyle instructions include eating 3 balanced meals daily, including breakfast; eating a diet high in fiber, vegetables, whole grains, lean protein, and healthy fats like olive oil and nuts; using the plate method for portion control[32] (Figure 1); avoiding sweetened beverages; daily walking or swimming; and adherence to medication. A registered dietitian is valuable in teaching and counseling the patient with diabetes regarding healthy eating and calorie goals.[33,34] Pedometers help patients monitor the number of steps taken each day and may motivate patients to increase activity.[35] Finally, practitioners should positively reinforce actual outcomes including weight loss, weight maintenance, and glycemic and cardiometabolic control.

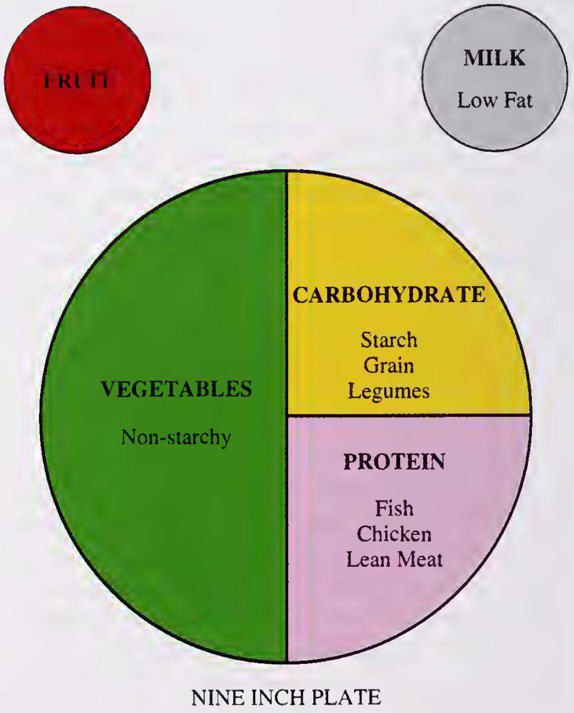


Fig 1. Plate method for portion control.

DIABETES MEDICATIONS: GLYCEMIC CONTROL, WEIGHT GAIN, AND HYPOGLYCEMIA

Large randomized controlled trials in both type 1 DM (T1DM) and T2DM demonstrate reduction of microvascular disease with intensive glycemic control, but at the price of excess weight gain and hypoglycemia. To date, it remains unclear whether intensive glycemic control improves cardiovascular outcomes. The United Kingdom Prospective Diabetes Study (UKPDS) was a 10-year trial of newly diagnosed T2DM patients randomized to intensive therapy with a sulfonylurea or insulin versus conventional therapy with diet modification. The intensive group achieved a lower median A_{1c} of 7.0% versus 7.9% in the conventional group, and a 25% reduction in microvascular complications.[36] All the groups accumulated weight over the 10-year study, but the intensive group gained 3.1 kg more than the conventional group. Weight gain was most pronounced in insulin users as compared with patients taking the sulfonylureas chlorpropramide and glibenclamide (4.0 kg, 2.6 kg, and 1.7 kg greater than conventional group, respectively).[36] A subgroup of overweight subjects treated with metformin achieved an A_{1c} of 7.4% versus 8.0% in the conventional group, and had a similar increase in weight as the conventional group but less than the groups treated with sulfonylureas or insulin.[37]

> *Large randomized controlled trials in both type 1 and type 2 diabetes demonstrate reduction of microvascular disease with intensive glycemic control, but at the price of excess weight gain and hypoglycemia.*

In the Diabetes Control and Complications Trial (DCCT), individuals with T1DM who received an intensive insulin regimen for 6.5 years and achieved an average A_{1c} of 7% had a decrease in retinopathy, nephropathy, and neuropathy.[38] The intensive group gained a mean of 4.6 kg more than the conventionally managed group, and the risk of hypoglycemia was 3× greater in the intensive group. Weight gain also worsened lipids and blood pressure.[39] Both the UKPDS and DCCT studies showed glycemic control, or achieving an A_{1c} of 7%, reduced microvascular complications, including retinopathy, nephropathy, and neuropathy.[36,38] However, the patients in the intensive groups were more vulnerable to weight gain and severe hypoglycemia.

Observational follow-up of the diabetic participants in UKPDS and DCCT for 10–11 years after trial completion demonstrated that, despite convergence of A_{1c}s between the intensive and conventional groups (at year 1 in T2DM and at year 5 in T1DM), individuals initially assigned to intensive control maintained persistence of microvascular benefit. Furthermore, there was an emergence of CV and mortality benefit in the intensive groups.[40,41] This suggests that glycemic control early on in the course of T1DM and T2DM may establish a foundation of vascular health or "metabolic programming" that protects against microvascular disease and possibly CVD and mortality later on in life.

Contrary to established benefits on microvascular disease, some studies have shown that intensive glycemic control compared with standard glycemic control may not improve CV risk,[42–44] and may increase mortality in specific populations with T2DM.[42] The Action to Control Cardiovascular Risk in Diabetes (ACCORD), Action in Diabetes and Vascular Disease (ADVANCE), and Veterans Affairs Diabetes Trial (VADT) studies randomized subjects with T2DM to intensive glycemic control (A_{1c}<6% to 6.5%) or standard glycemic control and found that intensive glycemic control did not significantly reduce CV events or mortality. Moreover, the ACCORD study was terminated early after 3.5 years due to increased mortality in those receiving intensive therapy. As compared with UKPDS subjects, those in the ACCORD, ADVANCE, and VADT trials were older (in their 60s, versus a median age of 54 years), had longer duration of DM (8–12 years versus newly diagnosed), and had a greater prevalence of CVD risk factors and CVD at baseline. Furthermore, in the intensively managed individuals in ACCORD, the rate of A_{1c} reduction was rapid (from 8.1% to 6.7% in 4 months), incidence of hypoglycemia requiring assistance was high (16% compared with 5% in conventional group; $P < 0.001$), and weight gain was clinically significant (greater than 10.0 kg weight gain in 28% versus 14% in the conventional group; $P < 0.001$). Thus, rapid lowering of blood glucose using multiple different medication classes led to more hypoglycemia, weight gain, and mortality in patients who had longstanding DM and background CVD.

Taking these studies into consideration, the American Diabetes Association currently recommends an A_{1c} goal less than 7%, with individualization of the goal based on patient characteristics.[33] Those with new-onset DM and low CV burden may be candidates for a lower A_{1c} goal if it can be achieved while simultaneously "avoiding significant hypoglycemia" and "other adverse effects of treatment"; the latter may be interpreted to include weight gain. The evidence directs providers to tailor their DM management and glycemic goals for individual patients, taking into account age/life expectancy, duration of DM, underlying CVD or advanced microvascular disease, risk of hypoglycemia, and ability to self-manage DM.[33] Regular, scheduled follow-up visits and judicious adjustment of DM medications based on A1c and glucose, may help avoid hypoglycemia and concomitant rapid weight gain, with hypoglycemia possibly being directly associated with mortality, and weight gain being a cophenomenon of hypoglycemia, which is associated with more complications in T2DM.

PHARMACOLOGIC MANAGEMENT OF TYPE 2 DIABETES: EFFECT ON GLYCEMIC CONTROL AND WEIGHT MANAGEMENT

Although lifestyle modification is a central component to DM management, many patients are unable to achieve optimal glycemic control without the addition of medication. The weight gain associated with certain DM medication in the UKPDS and other DM trials was confirmed in an observational real-world examination of 9456 patients in the Kaiser Permanente Northwest health system. Exposure to 1 year of sulfonylurea, insulin, TZDs, or metformin alone resulted in weight gain in all groups except metformin (+1.8 kg, +3.3 kg, +5.0 kg, −2.4 kg; $P < 0.05$ for all comparisons).[45] When initiating medications for glycemic management, providers and patients should be aware of medications' impact on weight and develop strategies to minimize this adverse side effect (Table 1).

Metformin is the first-line medication for management of type 2 diabetes.

Metformin is the first-line medication for management of T2DM.[46] It should be started upon diagnosis as long as there are no contraindications, including renal insufficiency, decompensated heart failure, liver disease, hypoxemic lung disease, and a history of lactic acidosis. It is especially preferred in overweight patients with T2DM, as it facilitates weight loss or weight maintenance,[47–55] reduces microvascular disease,[37,40] and may reduce CVD,[37,40] mortality,[37,40] and cancer.[56] Metformin decreases liver gluconeogenesis (a primary site of dysregulation in T2DM), increases peripheral glucose uptake, and reduces glucose absorption in the gut. It improves fasting glucose control, reduces A_{1c} by 1.0% to 2.0%, and has minimal risk of hypoglycemia.[46] Metformin has neutral to beneficial effects on lipids and BP, reducing total cholesterol by 5%.[57] Metformin is thought to facilitate weight loss by decreasing insulin levels, inducing anorexia or abdominal discomfort, and decreasing hunger,[58] possibly by increasing the gut hormone GLP−1.[58–60] Individuals may start at the lowest dose of metformin (500 mg once or twice daily with meals), self-titrate metformin to glucose goals, and use the maximum tolerated dose.[46] Patients should be cautioned about GI side effects, including loose stools and nausea; and if these occur, the metformin dose may be lowered to a tolerated dose or switched to a long-acting formulation (extended-release 500-mg or 750-mg tab with dinner). Given its potent glucose lowering, associated weight benefits, improved microvascular outcomes and possible CV outcomes in UKPDS, lack of hypoglycemia, and long-term safety data, practitioners should ensure that

Table 1. *Effect of Medications and Lifestyle Intervention on A_{1c} and Weight.*

Intervention	A_{1c} Reduction Expected, %	Weight Effect
Metformin[45–55]	1.0–2.0	Loss (0.6–2.7 kg)
GLP-1 analogs[46,96–100]	0.5–1.5	Loss (1.8–6.0 kg)
Pramlinitide[120–122]	0.5–0.7	Loss (1.5 kg)
AGIs[46,63–65]	0.5–0.8	Neutral
DPP-IV inhibitors[46,107,110–112]	0.5–0.8	Neutral
Sulfonylureas[45,46,70]	1.0–2.0	Gain (1.8–5.0 kg)
Meglitinides[46,72]	0.5–1.5	Gain (0.7–1.8 kg)
TZDs[46,69,85]	0.5–1.4	Gain (1.3–4.8 kg)
Insulin[36,46,75–77]	1.5–3.5	Gain (variable)
Intensive Lifestyle intervention[28]	0.6	Loss (8.6% of body weight)

Abbreviations: A_{1c}, glycated hemoglobin A_{1c}; AGIs, alpha-glucosidase inhibitors; DPP-IV, dipeptidyl peptidase-4; GLP-1, glucagon-like peptide-1; TZDs, thiazolidinediones.

metformin is a requisite component of the T2DM regimen and become very familiar with metformin prescribing and titration.

AGIs, including acarbose and miglitol, inhibit the alpha-glucosidase enzyme in the small intestine, reduce degradation of carbohydrates into monosaccharides, and diminish absorption of mealtime carbohydrates. Thus, AGIs improve postprandial glycemic control and lower A_{1c} by 0.5% to 0.8%.[46] They may also work by increasing GLP-1.[61] Side effects include flatulence, abdominal discomfort, and diarrhea due to fermentation of the nondigested carbohydrates in the colon.[39,62] AGIs induce weight loss of about 0.5–0.8 kg or are weight neutral,[63–65] and a meta-analysis found that monotherapy reduced BMI by 0.17 kg/m^2.[66] Acarbose is also associated with a reduced risk of CV events and myocardial infarction,[67] possibly related to reduction in postprandial glucose excursions, which are associated with oxidative stress and inflammation of the vasculature.[68]

Sulfonylureas stimulate insulin secretion from pancreatic beta cells by closing the adenosine triphosphate (ATP)-dependent potassium channels. Sulfonylureas potently lower A_{1c} by 1.0% to 2.0%,[46] but concomitantly increase the risk of hypoglycemia, induce modest weight gain, and have poor durability of effect over time.[69] Since sulfonylureas may stimulate insulin secretion even in the face of normal glucose levels, hypoglycemia is a common side effect, occurring in 28% of UKPDS sulfonylurea users as minor hypoglycemia and in 1.3% as major hypoglycemia requiring third-party intervention.[70] Weight gain is approximately 1.8 kg over 1 year and 5 kg over a 6-year period.[45,70] The weight gain may be attributed to effects of increased insulin levels on reduction of glycosuria and increased lipogenesis, and increased food intake to correct hypoglycemic events. Thus, when initiating sulfonylureas, one must start at a low dose and titrate slowly to avoid hypoglycemia and excess weight gain. Sulfonylureas should be used with caution in populations who are more vulnerable to hypoglycemia, including the elderly or frail, those with erratic eating patterns, and those with renal insufficiency. In a large retrospective cohort study of >90,000 users of oral diabetes agents in the United Kingdom, sulfonylurea monotherapy was associated with an increased risk of congestive heart failure and all-cause mortality as compared with metformin monotherapy.[71]

Meglitinides (repaglinide and nateglinide) are short-acting secretagogues that stimulate beta-cell insulin secretion by inhibiting the ATP-dependent potassium channels.[39] These agents are rapid- and short-acting as compared with long-acting sulfonylureas, are administered before meals to control postprandial glucose excursion, and lower A_{1c} by 0.5% to 1.5%.[46] Repaglinide is more potent than nateglinide, but also causes more weight gain and hypoglycemia (A_{1c} decrease of 1.57% versus 1.04%; weight gain of 1.8 kg versus 0.7 kg over 16 weeks, respectively).[72] Meglitinides are a good option for patients with variable eating patterns who require postprandial glycemic control.

Insulin injections are effective in achieving glucose control in T2DM when oral medications are not efficacious or are contraindicated. However, insulin is associated with weight gain that may in turn worsen insulin resistance, increase insulin requirements further, and perpetuate a vicious cycle of weight gain and worsening cardiometabolic profile. Weight gain is seen in both T1DM and T2DM using subcutaneous insulin.[36,38,39] In regard to basal insulins, the newer insulin analogs, detemir and glargine, deliver a long-lasting basal supply of insulin with relatively no peak, consistent insulin levels from dose to dose, and fewer hypoglycemic events as compared with neutral protamine Hagedorn (NPH) insulin.[73] NPH insulin has a peak effect at 6–8 hours after injection, more variability from dose to dose,[74] and more hypoglycemia and weight gain. As such, when choosing basal insulin, detemir and glargine may be superior in minimizing hypoglycemia and weight gain. Among the 2 basal insulins, detemir is associated with less weight gain[75–77]; this may be related to its albumin binding and preferential delivery to the liver as compared to the adipose. In more obese patients there appears to be weight loss with detemir[78]; however, higher doses of detemir insulin are required to achieve control, and twice-daily injections are required in about 50% of T2DM subjects to achieve 24-hour coverage.[75,77]

Weight gain is seen in both T1DM and T2DM using subcutaneous insulin.

Potential mechanisms of weight gain associated with insulin may include reduction of glycosuria, reactive consumption of sugary beverages and foods to avoid or treat hypoglycemia, differential effects of insulin analogs (detemir and glargine) and human insulin (regular and NPH insulin) on adipose tissue (lipogenesis) and liver (gluconeogenesis), effects on the central nervous system satiety center, and mode of current insulin delivery being subcutaneous as compared with the physiologic route being intraportal. Whenever possible, metformin should be used in conjunction with insulin in order to reduce insulin requirements.[51,52,79] If a patient requires escalating doses of basal insulin, ie, >0.8–1.0 units/kg/day of detemir or 0.5–0.7 units/kg/day of glargine,[75–77,80] prandial rapid-acting insulin such as aspart, lispro, or glulisine may be added to the largest meal of the day or to each meal. Using very high basal insulin doses to cover both basal and carbohydrate metabolism may lead to a vicious cycle of eating to avoid hypoglycemia and will be prohibitive to a patient who is trying to make diet modifications and lose weight.

The TZDs, rosiglitazone and pioglitazone, improve glycemic control by increasing insulin-mediated glucose uptake in the liver, adipose tissue, and skeletal muscle. They activate peroxisome proliferator-activated receptor gamma (PPARγ), which are nuclear transcription factors that enhance glucose transporter type 4 (GLUT-4) expression.[81–83] TZDs may also increase adiponectin synthesis and secretion, which has been associated with improved insulin action.[84] TZDs lower A_{1c} by 0.5% to 1.4%,[46] induce weight gain, but do not increase hypoglycemia when used alone or with metformin. The ADOPT (A Diabetes Outcome and Progression Trial) study found rosiglitazone to have a more durable effect on glucose-lowering over 5 years of monotherapy, as compared with metformin and glyburide. However, rosiglitazone users also had greater weight gain, of 4.8 kg, compared with 1.6 kg in glyburide users and a weight loss of 2.9 kg in the metformin group.[69] Higher doses of TZDs induce slightly more A_{1c} lowering, but this is also paralleled by greater weight gain. In a randomized controlled trial, A_{1c} was reduced 1.0%, 0.9%, and 1.4% (for pioglitazone 15 mg, 30 mg, and 45 mg daily, respectively), but weight gain from baseline in each group was 1.3 kg, 1.3 kg, and 2.8 kg, respectively.[85] Pioglitazone is associated with improved total cholesterol, TG, and low-density lipoprotein (LDL) levels,[86] and both TZDs are associated with systolic BP reduction of 4–5 mm Hg and diastolic BP reduction of 2–4 mm Hg.[87] The homeostasis model of assessment showed improved beta-cell function and decreased insulin resistance with rosiglitazone.[88] Improved insulin sensitivity in the setting of concomitant weight gain with TZDs may be explained by a decrease in visceral fat, as shown indirectly by reduced waist-to-hip ratio,[88] and an increase in subcutaneous fat. Visceral fat harbors inflammatory cells and produces inflammatory mediators that reduce insulin sensitivity.[82] Some of the weight gain is also due to fluid retention attributed to increased vascular permeability factors, enhanced endothelial-mediated vasodilation,[82] and sodium retention by the kidney. Thus, TZDs increase the risk for heart failure.[89] Despite beneficial effects on insulin sensitivity, there is conflicting data regarding rosiglitazone's association with increased cardiovascular events.[71,90,91] Currently, pioglitazone is preferred if choosing TZD therapy.[46,71] Providers should be aware of the increased risk of fractures with TZDs.[92,93]

The GLP-1 analogs are a newer class of injectable medications that mimic gut hormone action, and include exenatide and liraglutide. In normal humans, an oral glucose load triggers the enteroendocrine axis to release gut hormones from the distal intestine. GLP-1 and glucose-dependent insulinotropic peptide (GIP) enhance insulin secretion in response to glucose.[94] Individuals with T2DM exhibit abnormal secretion of GLP-1 and diminished response to GLP-1.[95] GLP-1 reduces postprandial glucose by enhancing glucose-dependent insulin secretion, suppresses glucagon secretion, slows gastric emptying, and curbs appetite presumably through central nervous system effects. Exenatide and liraglutide are synthetic analogues of GLP-1 that are resistant to degradation by DPP-IV. Trials evaluating exenatide as add-on therapy to other oral medications (metformin, sulfonylureas, combined metformin and sulfonyureas, TZDs) and as monotherapy show reduced A_{1c} by 0.5% to 1% and dose-dependent weight loss of 1.8–2.8 kg after 12–30 weeks,[96–100] with greater weight loss when used in conjunction with metformin as compared with other agents. A selected group of individuals from these trials who responded to and tolerated exenatide and continued it for 3 years showed progressive weight loss over time of up to 6 kg, durability of A_{1c} effect with slight deterioration toward the end of 3 years, and significant reduction in BP and improvements in all lipid parameters.[101] The greatest percentage change in A_{1c} due to exenatide appears to occur at 12 weeks, but weight loss may continue.[96–100] If there is no improvement in glucose or weight after 3 months of treatment with exenatide, providers should consider discontinuing the medication. Exenatide largely controls postprandial glucose excursions, and as such may reduce inflammation and cardiovascular events; however, at this time there is no evidence for complication reduction. Since GLP-1 and exenatide have been shown to promote beta-cell proliferation and neogenesis and reduce apoptosis in rodents and cell lines,[94]

> *GLP-1 analogs, including exenatide and liraglutide, are a newer class of injectable medications that mimic gut hormone action.*

there is hope that it may preserve beta-cell function in humans. Liraglutide is a once-daily GLP-1 analog recently approved by the US Food and Drug Administration for management of T2DM. Liraglutide reduces body weight (1.8–2.8 kg depending on dose) and A_{1c} by 0.6% to 1.5%.[99,102–106] Nausea tends to be the main side effect limiting use and titration of the GLP-1 analogs. They are contraindicated in patients with gastroparesis because of the delay in gastric emptying.[107] When adding a GLP-1 analog to a sulfonylurea, the sulfonylurea dose may need to be reduced to minimize hypoglycemia. Both GLP-1 analogs are associated with an increased risk of pancreatitis,[108] and liraglutide is implicated as a carcinogen in rodents, causing increased rates of medullary thyroid carcinoma.[109]

DPP-IV inhibitors inhibit the peptidase that degrades endogenous GLP-1, thus raising levels of GLP-1 to approximately twice normal. Sitagliptin and saxagliptin are oral DPP-IV inhibitors that improve A_{1c} modestly by 0.5% to 0.8% and are weight neutral.[46,107,110–112] These medications may be useful in individuals who are nearing the goal A_{1c} and wish to avoid weight gain, and in individuals who are at risk for hypoglycemia (the elderly, those with renal insufficiency, and frail individuals). In a meta-analysis, sitagliptin was associated with increased headache and subtle immune system effects,[113] including nasopharyngitis and urinary tract infection.[114]

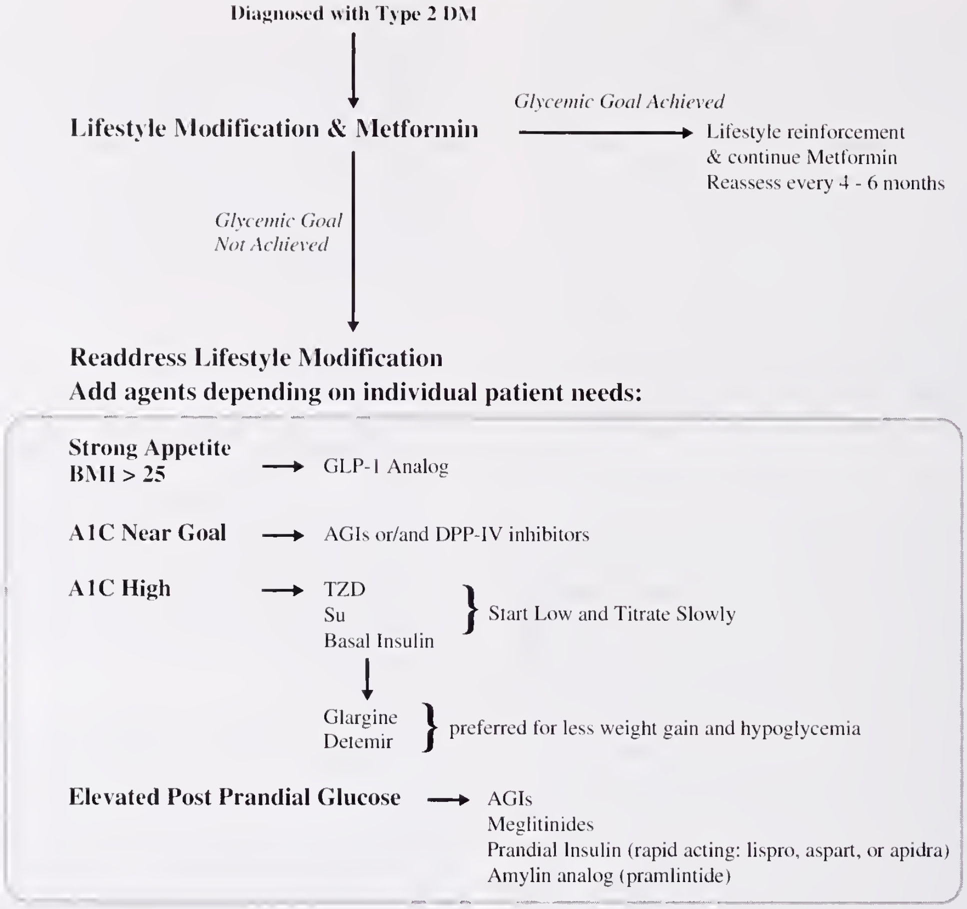
Pramlintide is an amylin analogue approved for use in T1DM and T2DM, as an adjunct to mealtime insulin injections, to control postprandial glucose excursions. Amylin is normally cosecreted with insulin from pancreatic beta cells, but this response is absent or reduced in T1DM and T2DM.[115,116] Pramlinitide suppresses postprandial glucagon secretion, slows gastric emptying, and suppresses appetite.[117–119] A single dose of premeal pramlintide reduced food intake by 16% to 23%.[119] Pramlinitide improves glycemic control, with an A_{1c} reduction of 0.5% to 0.7%, and induces weight maintenance or weight loss of approximately 1.5 kg in patients with T2DM.[120–122] It has a similar but less-potent effect on glucose lowering and weight loss in T1DM. The limiting side effects include nausea, vomiting, constipation, and hypoglycemia, with the latter being minimized by appropriate reduction in mealtime insulin (Table 1).

DIABETES MELLITUS MANAGEMENT STRATEGIES

Management of T2DM requires equal attention to weight management and glycemic control. Lifestyle modification is the cornerstone of T2DM management, and healthy behaviors should be reinforced on a regular basis. At each encounter, providers should assess for change and improvement in patient behavior, weight, glucose, and other metabolic parameters. Patients should be given discrete, simple instructions for lifestyle change, including portion control using the plate method, avoiding sweetened beverages, and participating in 30 minutes of moderate physical activity 5 days a week (Figure 2 and Figure 3).

> *Patients should be given discrete, simple instructions for lifestyle change, including portion control using the plate method, avoiding sweetened beverages, and participating in 30 minutes of moderate physical activity 5 days a week.*

Metformin is the first-line medication initiated upon diagnosis in patients with T2DM, given its potent glucose-lowering and associated weight benefits. It may be used in combination with all other diabetes medications, including insulin. For individuals being treated with other oral medications, metformin may be added to the regimen or may replace the oral medication, depending on baseline glucose control. In overweight patients who are unable to achieve glycemic goals on maximum metformin therapy and lifestyle modification, and who may have a voracious appetite that drives overeating, injectable GLP-1 agonists (exenatide, liraglutide) may be added to facilitate weight loss, normalize satiety, and improve postprandial glucose excursions. In individuals with an A_{1c} that is approaching goal, weight-neutral oral medications with low rates of hypoglycemia include the AGIs (acarbose, miglitol) used at meals and the DPP-IV inhibitors (sitagliptin, saxagliptin) used once daily. If using an insulin secretagogue or TZD for potency of A_{1c} lowering, the risk for hypoglycemia with the former class, and weight gain with both classes, should be kept in mind. Sulfonylureas should be initiated at a low dose and titrated gradually based on individual response in order to minimize hypoglycemia and concomitant weight gain. Though initially potent, sulfonylureas have poor durability over time. Individuals at risk for hypoglycemia with sulfonylureas, ie. those who are elderly, frail, have renal insufficiency, or have irregular eating habits, may benefit from rapid-acting meglitinides


Diagnosed with Type 2 DM
Lifestyle Modification & Metformin
Glycemic Goal Achieved
Lifestyle reinforcement
& continue Metformin
Reassess every 4 - 6 months
Glycemic Goal
Not Achieved
Readdress Lifestyle Modification
Add agents depending on individual patient needs:
Strong Appetite
BMI > 25
GLP-1 Analog
A1C Near Goal
AGIs or/and DPP-IV inhibitors
A1C High
TZD
Su
Basal Insulin
Start Low and Titrate Slowly
Glargine
Detemir
preferred for less weight gain and hypoglycemia
Elevated Post Prandial Glucose
AGIs
Meglitinides
Prandial Insulin (rapid acting: lispro, aspart, or apidra)
Amylin analog (pramlintide)

GLP-1 Analog = Glucagon-like peptide-1 Analog
DPP-IV = Dipeptidyl peptidase-IV inhibitors
Su = Sulfonylurea
AGIs = Alpha-Glucosidase Inhibitors
TZD = Thiazolidinediones

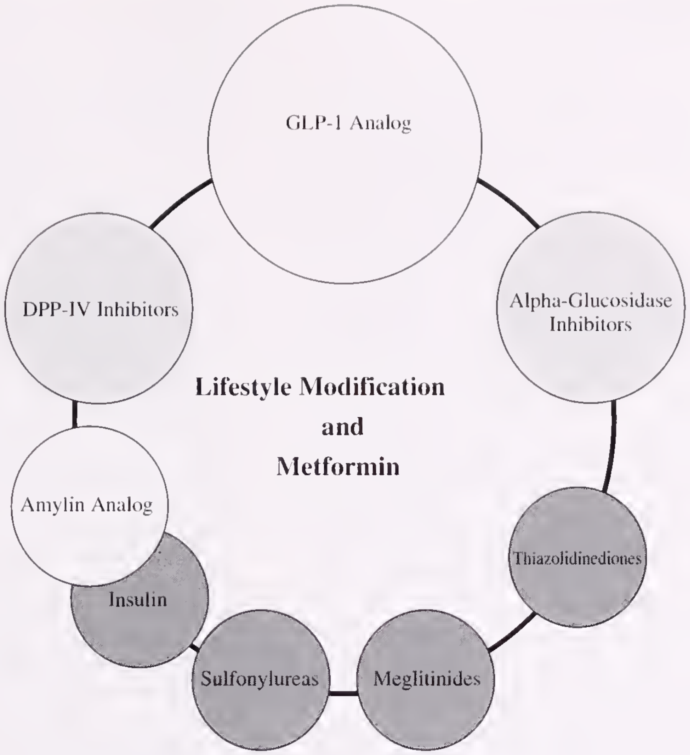
This is a guideline for glycemic management with a focus on weight control.
As always, therapeutic choices are guided by clinical decision making, including patient characteristics, contraindications to therapy, and interactions between medications.

Fig 2. Weight management guided algorithm for outpatient type 2 diabetes mellitus control.

(repaglinide, nateglinide) at mealtimes. This class also carries a risk, albeit a smaller risk, of hypoglycemia and weight gain, so should be titrated slowly. Of the 2 TZDs, pioglitazone is preferred, and the lower doses should be used to prevent excessive weight gain. The highest dose may not add much more glycemic effect, but will increase weight and fluid retention.

Many patients will not achieve adequate glycemic control with lifestyle modification and oral medications, and therefore must be started on insulin. When choosing basal insulin, providers may consider utilizing detemir and glargine over NPH, since the former insulin analogs have a more consistent profile from day to day, fewer hypoglycemic events, and less weight gain. In patients requiring escalating doses of basal insulin, mealtime rapid-acting insulin may be added to cover carbohydrate intake, and to avoid hypoglycemia and reactive overeating required to "keep up" with high basal insulin doses. Metformin may be used in combination with insulin to minimize weight gain and insulin requirements. Addition of pramlinitide injections to mealtime insulin may lower insulin requirements, improve postprandial glucose excursions, and reduce weight (Figure 2 and Figure 3).

The timetable to obtaining glycemic control is another important consideration. If reaching glucose goals is carried out too zealously and rapidly, it may contribute to severe hypoglycemia and inordinate weight gain, as seen in the ACCORD study.



1. Lifestyle modification and metformin are central to type 2 DM management.
   Reinforce and assess for changes in lifestyle at each visit. Titrate metformin to maximum tolerated dose.
2. If weight loss is a priority, use GLP-1 Analog (exenatide injections before meals twice daily or liraglutide once daily).
3. Alpha-glucosidase inhibitors and DPP-IV inhibitors are weight neutral and modestly reduce A1C.
4. If thiazolidinediones (pioglitazone) are used, minimize weight gain by utilizing lowest dose possible.
5. When treating with sulfonylureas or insulin, titrate slowly in order to avoid hypoglycemia and reactive eating which leads to weight gain.
6. When choosing long acting insulin, use detemir and glargine over NPH for less weight gain and hypoglycemia.
   If possible, continue metformin in insulin users to reduce insulin requirements and weight gain.
7. The amylin anolog, pramlintide, is injected before meals, reduces insulin requirements, and induces weight loss.

○ = medication associated with weight loss ◍ = weight neutral medication ● = medication associated with weight gain

Fig 3. Strategy for managing type 2 diabetes mellitus with awareness of weight side effects.

If reaching glucose goals is carried out too zealously and rapidly, it may contribute to severe hypoglycemia and inordinate weight gain.

A_{1c} and glucose goals are important, but achieving those goals safely is probably even more important, particularly in vulnerable populations. In addition, if the patient reduces caloric intake, increases physical activity, and/or loses weight, the provider and patient must be ready to reduce doses of medication including insulin and insulin secretagogues to prevent hypoglycemic episodes. It is important to assess for hypoglycemia at each patient encounter, educate the patient about hypoglycemia, and titrate medications appropriately at time intervals of approximately 4–12 weeks. Gradual titration of glucose-lowering medications may minimize hypoglycemia and/or weight gain while allowing the individual to incorporate lifestyle changes that may improve insulin sensitivity and reduce medication requirements.

CONCLUSION

Obesity and the accompanying metabolic syndrome are independent risk factors for CVD and mortality in T2DM. Weight loss ameliorates the dysmetabolic state and associated CV risk factors, including BP, blood glucose, lipids, and inflammatory markers. Although tight glycemic control definitively reduces microvascular events, including retinopathy, nephropathy, and neuropathy, the same cannot yet be concluded for glycemic control and CV events and mortality. Unfortunately, individuals with T2DM who reach glycemic control often carry the burden of abnormal weight gain attributable to various factors, including certain medication classes and hypoglycemia. The introduction of newer classes and formulations of diabetes medications provides opportunities to control weight and the associated dysmetabolic state. This may translate into long-term effects on CVD and mortality. In order to face the clinical challenge of managing patients with T2DM, providers should familiarize themselves with therapeutic strategies that have the least-detrimental effects on patient weight while improving glycemic control and overall outcomes.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Morris RD, Rimm DL, Hartz AJ, et al. Obesity and heredity in the etiology of non-insulin dependent diabetes mellitus in 32,662 adult white women. *Am J Epidemol* 1989; 130: 112–121.
2. Skarfors ET, Selinus KI, Lithell HO. Risk factors for developing non-insulin dependent diabetes: a 10 year follow up of men in Uppsala. *BMJ* 1991; 303: 755–760.
3. Charles MA, Fontbonne A, Thibult N, et al. Risk factors for NIDDM in white population: Paris prospective study. *Diabetes* 1991; 40: 796–799.
4. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus: provisional report of a WHO consultation. *Diabet Med* 1998; 15: 539–553.
5. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Executive Summary of the third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA* 2001; 285: 2486–2497.
6. Alberti KG, Zimmet P, Shaw J; IDF Epidemiology Task Force Consensus Group. The metabolic syndrome–a new worldwide definition. *Lancet* 2005; 366: 1059–1062.
7. Anderson JW, Konz EC. Obesity and disease management: effects of weight loss on comorbid conditions. *Obes Res* 2001; 9: 326S–334S.
8. Fox CS, Pencina MJ, Wilson PW, et al. Lifetime risk of cardiovascular disease among individuals with and without diabetes stratified by obesity status in the Framingham heart study. *Diabetes Care* 2008; 31: 1582–1584.
9. Eeg-Olofsson K, Cederholm J, Nilsson PM, et al. Risk of cardiovascular disease and mortality in overweight and obese patients with type 2 diabetes: an observational study in 13,087 patients. *Diabetologia* 2009; 52: 65–73.
10. Bonora E, Targher G, Formentini F, et al. The metabolic syndrome is an independent predictor of cardiovascular disease in type 2 diabetic subjects. Prospective data from the Verona Diabetes Complications Study. *Diabet Med* 2004; 21: 52–58.
11. Bonadonna RC, Cucinotta D, Dedele D, et al; for Metascreen Writing Committee. The metabolic syndrome is a risk indicator of microvascular and macrovascular complications in diabetes: results from Metascreen, a multicenter diabetes clinic-based survey. *Diabetes Care* 2006; 29: 2701–2707.
12. Knowler WC, Barrett-Connor E, Fowler SE, et al; Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Eng J Med* 2002; 346: 393–403.
13. Goldberg RB, Temprosa M, Haffner S, et al. Effect of progression from impaired glucose tolerance to diabetes on cardiovascular risk factors and its amelioration by lifestyle and metformin intervention: the Diabetes Prevention Program randomized trial by the Diabetes Prevention Program Research Group. *Diabetes Care* 2009; 32: 726–732.
14. Williams KV, Kelley DE. Metabolic consequences of weight loss on glucose metabolism and insulin action in type 2 diabetes. *Diabetes Obes Metab* 2000; 2: 121–129.
15. Klein S, Sheard NF, Pi-Sunyer X, et al. Weight management through lifestyle modification for the prevention and management of type 2 diabetes: rationale and strategies. *Diabetes Care* 2004; 27: 2067–2073.
16. Anderson JW, Kendall CW, Jenkins DJ. Importance of weight management in type 2 diabetes: review with meta-analysis of clinical studies. *J Am Coll Nutr* 2003; 22: 331–339.
17. Norris SL, Zhang X, Avenell A, et al. Long term effectiveness of lifestyle and behavioral weight loss interventions in adults with type 2 diabetes: a meta-analysis. *Am J Med* 2004; 117: 762–774.
18. Ryan M, Livingstone MB, Ducluzeau PH, et al. Is a failure to recognize an increase in food intake a key to understanding insulin-induced weight gain? *Diabetes Care* 2008; 31: 448–450.
19. Close EJ, Wiles PG, Lockton JA, et al. The degree of day-to day variation in food intake in diabetic patients. *Diabet Med* 1993; 10: 514–520.
20. Ford ES, Herman WH. Leisure-time physical activity patterns in the U.S. diabetic population. *Diabetes Care* 1995; 18: 27–33.

21. Hays LM, Clark DO. Correlates of physical activity in a sample of older adults with type 2 diabetes. *Diabetes Care* 1999; 22: 706–712.
22. Sigal RJ, Kenny GP, Boulé; NG, et al. Effects of aerobic training, resistance training or both on glycemic control in type 2 diabetes: a randomized trial. *Ann Intern Med* 2007; 147: 357–369.
23. Boulé; NG, Haddad E, Kenny GP, et al. Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus: a meta-analysis of controlled clinical trials. *JAMA* 2001; 286: 1218–1227.
24. Payne WR, Walsh KJ, Harvey JT, et al. Effect of a low-resource-intensive lifestyle modification program incorporating gymnasium-based and home-based resistance training on type 2 diabetes risk in Australian adults. *Diabetes Care* 2008; 31: 2244–2250.
25. Praet SF, van Rooij ES, Wijtvliet A, et al. Brisk walking compared with an individualised medical fitness programme for patients with type 2 diabetes: a randomised controlled trial. *Diabetologia* 2008; 51: 736–746.
26. Guare JC, Wing RR, Grant A. Comparison of obese NIDDM and nondiabetic women: short-and long-term weight loss. *Obesity Res* 1995; 3: 329–335.
27. Bitz C, Toubro S, Larsen TM, et al. Increased 24-h energy expenditure in type 2 diabetes. *Diabetes Care* 2004; 27: 2416–2421.
28. Pi-Sunyer X, Blackburn G, Brancati FL, et al.; for the Look AHEAD Research Group. Reduction in weight and cardiovascular disease risk factors in individuals with type 2 diabetes: one-year results of the Look AHEAD trial. *Diabetes Care* 2007; 30: 1374–1383.
29. Redmon JB, Bertoni AG, Connelly S, et al.; for the Look AHEAD Research Group. Effect of the Look AHEAD study intervention on medication use and related cost to treat cardiovascular disease risk factors in individuals with type 2 diabetes. *Diabetes Care* 2010; 33: 1153–1158.
30. Norris SL, Lau J, Smith SJ, et al. Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. *Diabetes Care* 2002; 25: 1159–1171.
31. Wishnofsky M. Caloric equivalents of gained or lost weight. *Am J Clin Nutr* 1958; 6: 542–546.
32. Camelon KM, Hadell K, Jämsén PT, et al.; DAIS Project Group. The Plate Model: a visual method of teaching meal planning. Diabetes Atherosclerosis Intervention Study. *J Am Diet Assoc* 1998; 98: 1155–1158.
33. American Diabetes Association. Standards of Medical Care in Diabetes–2010 [published correction appears in *Diabetes Care* 2010;33:692]. *Diabetes Care* 2010; 33(suppl 1): S11–S61.
34. American Diabetes Association. Nutrition recommendations and interventions for diabetes: a position statement of the American Diabetes Association. *Diabetes Care* 2008; 31(suppl 1): S61–S78.
35. Bravata DM, Smith-Spangler C, Sundaram V, et al. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA* 2007; 298: 2296–2304.
36. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood glucose control with sulphonylureas, or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33) [published correction appears in *Lancet* 1999;354:602]. *Lancet* 1998; 352: 837–853.
37. UK Prospective Diabetes Study (UKPDS) Group. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34) [published correction appears in *Lancet* 1998;352:1558]. *Lancet* 1998; 352: 854–865.
38. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development of and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 1993; 329: 977–986.
39. Purnell JQ, Hokanson JE, Marcovina SM, et al. Effect of excessive weight gain with intensive therapy of type 1 diabetes on lipid levels and blood pressure: results from the DCCT. Diabetes Control and Complications Trial [published correction appears in *JAMA* 1998;280:1484]. *JAMA* 1998; 280: 140–146.
40. Holman RR, Paul SK, Bethel MA, et al. 10-year follow up of intensive glucose control in type 2 diabetes. *N Engl J Med* 2008; 359: 1577–1589.
41. The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) Study Research Group. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med* 2005; 353: 2643–2653.
42. Gerstein HC, Miller ME, Byington RP, et al.; Action to Control Cardiovascular Risk in Diabetes Study Group. Effects of intensive glucose lowering in type 2 diabetes. *N Engl J Med* 2008; 358: 2545–2559.
43. Patel A, MacMahon S, Chalmers J, et al.; ADVANCE Collaborative Group. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *N Engl J Med* 2008; 358: 2560–2572.
44. Duckworth W, Abraira C, Mortiz T, et al.; VADT Investigators. Glucose control and vascular complications in veterans with type 2 diabetes [published corrections appear in *N Engl J Med* 2009; 361: 1028 and 2009;361:1024–1025]. *N Engl J Med* 2009; 360: 129–139.
45. Nichols GA, Gomez-Caminero A. Weight changes following the initiation of new anti-hyperglycaemic therapies. *Diabetes Obes Metab* 2007; 9: 96–102.
46. Nathan DM, Buse JB, Davidson MB, et al. Medical management of hyperglycemia in type 2 diabetes mellitus: a consensus algorithm for the initiation and adjustment of therapy: a consensus statement from the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care* 2009; 32: 193–203.
47. Purnell JQ, Weyer C. Weight effect of current and experimental drugs for diabetes mellitus: from promotion to alleviation of obesity. *Treat Endocrinol* 2003; 2: 33–47.
48. Johansen K. Efficacy of metformin in the treatment of NIDDM: meta-analysis. *Diabetes Care* 1999; 22: 33–37.
49. Inzucchi SE, Maggs DG, Spollett GR, et al. Efficacy and metabolic effects of metformin and troglitazone in type II diabetes mellitus. *N Engl J Med* 1998; 338: 867–872.
50. DeFronzo RA, Goodman AM. Efficacy of metformin in patients with NIDDM. *N Engl J Med* 1995; 333: 541–549.

51. Relimpio F, Pumar A, Losada F, et al. Adding metformin versus insulin dose increase in insulin-treated but poorly controlled type 2 diabetes mellitus: an open-label randomized trial. *Diabet Med* 1998; 15: 997–1002.
52. Avilés-Santa L, Sindig J, Raskin P. Effects of metformin in patients with poorly controlled, insulin-treated type 2 diabetes mellitus: a randomized, double-blind, placebo-controlled trial. *Ann Intern Med* 1999; 131: 182–188.
53. Stumvoll M, Nurjhan N, Perriello G, et al. Metabolic effects of metformin in non-insulin-dependent diabetes mellitus. *N Engl J Med* 1995; 333: 550–554.
54. Hermann LS, Scherstén B, Bitzén PO, et al. Therapeutic comparison of metformin and sulfonylurea, alone and in various combinations: a double-blind controlled study. *Diabetes Care* 1994; 17: 1100–1109.
55. Dornan TL, Heller SR, Peck GM, Tattersall RB. Double-blind evaluation of efficacy and tolerability of metformin in NIDDM. *Diabetes Care* 1991; 14: 342–344.
56. Evans JM, Donnelly LA, Emslie-Smith AM, et al. Metformin and reduced risk of cancer in diabetic patients. *BMJ* 2005; 330: 1304–1305.
57. Wulffelé MG, Kooy A, de Zeeuw D, et al. The effect of metformin on blood pressure, plasma cholesterol and triglycerides in type 2 diabetes mellitus: a systematic review. *J Intern Med* 2004; 256: 1–14.
58. Lee A, Morley JE. Metformin decreases food consumption and induces weight loss in subjects with obesity with type II non-insulin-dependent diabetes. *Obes Res* 1998; 6: 47–53.
59. Green BD, Irwin N, Duffy NA, et al. Inhibition of dipeptidyl peptidase-IV activity by metformin enhances the antidiabetic effects of glucagon-like peptide-1. *Eur J Pharmacol* 2006; 547: 192–199.
60. Hinke SA, Kühn-Wache K, Hoffmann T, et al. Metformin effects on dipeptidylpeptidase IV degradation of glucagon-like peptide-1. *Biochem Biophys Res Commun* 2002; 291: 1302–1308.
61. Lee A, Patrick P, Wishart J, et al. The effects of miglitol on glucagon-like peptide-1 secretion and appetite sensations in obese type 2 diabetics. *Diabetes Obes Metab* 2002; 329–335.
62. Krentz AJ, Bailey CJ. Oral antidiabetic agents: current role in type 2 diabetes mellitus. *Drugs* 2005; 65: 385–411.
63. Hoffmann J, Spengler M. 24-week monotherapy with acarbose, glibenclamide, or placebo in NIDDM patents. The Essen study. *Diabetes Care* 1994; 17: 561–566.
64. Hoffman J, Spengler M. Efficacy of 24-week monotherapy with acarbose, metformin, or placebo in dietary-treated NIDDM patients: the Essen-II study. *Am J Med* 1997; 103: 483–490.
65. Wolever TM, Chiasson JL, Josse RG, et al. Small weight loss on long-term acarbose therapy with no change in dietary pattern or nutrient intake of individuals with non-insulin-dependent diabetes. *Int J Obes Relat Metab Disord* 1997; 21: 756–763.
66. Van de Laar FA, Lucassen PL, Akkermans RP, et al. Alpha-glucosidase inhibitors for patients with type 2 diabetes: results from a Cochrane systematic review and meta-analysis. *Diabetes Care* 2005; 28: 154–163.
67. Hanefeld M, Cagatay M, Petrowitsch T, et al. Acarbose reduces risk for myocardial infarction in type 2 diabetic patients: meta-analysis of seven long-term studies. *Eur Heart J* 2004; 25: 10–16.
68. Ceriello A, Hanefeld M, Leiter L, et al. Postprandial glucose regulation and diabetic complications. *Arch Intern Med* 2004; 164: 2090–2095.
69. Kahn SE, Haffner SM, Heise MA, et al.; for the ADOPT Study Group. Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *N Engl J Med* 2006; 355: 2427–2444.
70. UKPDS Group. United Kingdom Prospective Diabetes Study 24: a 6-year, randomized, controlled trial comparing sulfonylurea, insulin, and metformin therapy in patients with newly diagnosed type 2 diabetes that could not be controlled with diet therapy. *Ann Intern Med* 1998; 128: 165–175.
71. Tzoulaki I, Molokhia M, Curcin V, et al. Risk of cardiovascular disease and all-cause mortality among patients with type 2 diabetes prescribed oral antidiabetes drugs: retrospective cohort study using UK general practice research database. *BMJ* 2009; 339: b4731.
72. Rosenstock J, Hassman DR, Madder RD, et al. Repaglinide versus nateglinide monotherapy: a randomized, multicenter study. *Diabetes Care* 2004; 27: 1265–1270.
73. Horvath K, Jeitler K, Berghold A, et al. Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2007; 2: CD005613.
74. Heise T, Nosek L, Ronn BB, et al. Lower within-subject variability of insulin detemir in comparison to NPH insulin and insulin glargine in people with type 1 diabetes. *Diabetes* 2004; 53: 1614–1620.
75. Rosenstock J, Davies M, Home PD, et al. A randomised, 52-week, treat-to-target trial comparing insulin detemir with insulin glargine when administered as add-on to glucose-lowering drugs in insulin-naive people with type 2 diabetes. *Diabetologia* 2008; 51: 408–416.
76. Hollander P, Cooper J, Bregnhoj J, Pedersen CB. A 52-week, multinational, open-label, parallel-group, noninferiority, treat-to-target trial comparing insulin detemir with insulin glargine in a basal-bolus regimen with mealtime insulin aspart in patients with type 2 diabetes. *Clin Ther* 2008; 30: 1976–1987.
77. Swinnen SG, Dain MP, Aronson R, et al. A 24-week, randomized, treat-to-target trial comparing initiation of insulin glargine once-daily with insulin detemir twice-daily in patients with type 2 diabetes inadequately controlled on oral glucose-lowering drugs. *Diabetes Care* 2010; 33: 1176–1178.
78. Raslová K, Tamer SC, Clauson P, Karl D. Insulin detemir results in less weight gain than NPH insulin when used in basal-bolus therapy for type 2 diabetes mellitus, and this advantage increases with baseline body mass index. *Clin Drug Investig* 2007; 27: 279–285.
79. Wulffelé MG, Kooy A, Lehert P, et al. Combination of insulin and metformin in the treatment of type 2 diabetes. *Diabetes Care* 2002; 25: 2133–2140.
80. Holman RR, Farmer AJ, Davies MJ, et al. Three-year efficacy of complex insulin regimens in type 2 diabetes. *N Engl J Med* 2009; 361: 1736–1747.

81. Mitri J, Hamdy O. Diabetes medications and body weight. *Expert Opin Drug Saf* 2009; 8: 573–584.
82. Vasudevan AR, Balasubramanyam A. Thiazolidinediones: a review of their mechanisms of insulin sensitization, therapeutic potential, clinical efficacy and tolerability. *Diabetes Technol Ther* 2004; 6: 850–863.
83. Yki-Järvinen H. Thiazolidinediones. *N Engl J Med* 2004; 351: 1106–1118.
84. Phillips SA, Ciaraldi TP, Kong AP, et al. Modulation of circulating and adipose tissue adiponectin levels by antidiabetic therapy. *Diabetes* 2003; 52: 667–674.
85. Aronoff S, Rosenblatt S, Braithwaite S, et al. Pioglitazone hydrochloride monotherapy improves glycemic control in the treatment of patients with type 2 diabetes: a 6-month randomized placebo-controlled dose-response study. *Diabetes Care* 2000; 23: 1605–1611.
86. Van Wijk JP, de Koning EJ, Martens EP, Rabelink TJ. Thiazolidinediones and blood lipids in type 2 diabetes. *Arterioscler Thromb Vasc Biol* 2003; 23: 1744–1749.
87. Giles TD, Sander GE. Effects of thiazolidinediones on blood pressure. *Curr Hypertens Rep* 2007; 9: 332–337.
88. Lebovitz HE, Dole JF, Patwardhan R, et al. Rosiglitazone monotherapy is effective in patients with type 2 diabetes. *J Clin Endocrinol Metab* 2001; 86: 280–288.
89. Komajda M, McMurray JJ, Beck-Nielsen H, et al. Heart failure events with rosiglitazone in type 2 diabetes: data from the RECORD clinical trial. *Eur Heart J* 2010; 31: 824–831.
90. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes [published correction appears in *N Engl J Med* 2007;357:100]. *N Engl J Med* 2007; 356: 2457–2471.
91. Diamond GA, Bax L, Kaul S. Uncertain effects of rosiglitazone on the risk for myocardial infarction and cardiovascular death. *Ann Intern Med* 2007; 147: 578–581.
92. Douglas IJ, Evans SJ, Pocock S, Smeeth L. The risk of fractures associated with thiazolidinediones: a self-controlled case-series study. *PLoS Med* 2009; 6: e1000154.
93. Dormuth CR, Carney G, Carleton B, et al. Thiazolidinediones and fractures in men and women. *Arch Intern Med* 2009; 169: 1395–1402.
94. Drucker DJ. The biology of incretin hormones. *Cell Metab* 2006; 3: 153–165.
95. Vilsbøll T, Krarup T, Deacon CF, et al. Reduced postprandial concentrations of intact biologically active glucagon-like peptide-1 in type 2 diabetic patients. *Diabetes* 2001; 50: 609–613.
96. Buse JB, Henry RR, Han J, et al. Effects of exenatide (exendin-4) on glycemic control over 30 weeks in sulfonylurea-treated patients with type 2 diabetes. *Diabetes Care* 2004; 27: 2628–2635.
97. Kendall DM, Riddle MC, Rosenstock J, et al. Effects of exenatide (exendin-4) on glycemic control over 30 weeks in patients with type 2 diabetes treated with metformin and a sulfonylurea. *Diabetes Care* 2005; 28: 1083–1091.
98. DeFronzo RA, Ratner RE, Han J, et al. Effects of exenatide (exendin-4) on glycemic control and weight over 30 weeks in metformin-treated patients with type 2 diabetes. *Diabetes Care* 2005; 28: 1092–1100.
99. Zinman B, Hoogwerf BJ, Durán García S, et al. The effect of adding exenatide to a thiazolidinedione in suboptimally controlled type 2 diabetes: a randomized trial [published correction appears in *Ann Intern Med* 2007;146:896]. *Ann Intern Med* 2007; 146: 477–485.
100. Moretto TJ, Milton DR, Ridge TD, et al. Efficacy and tolerability of exenatide monotherapy over 24 weeks in antidiabetic drug–naive patients with type 2 diabetes: a randomized, double-blind, placebo-controlled, parallel-group study [published correction appears in *Clin Ther* 2008;30:1937]. *Clin Ther* 2008; 30: 1448–1460.
101. Klonoff DC, Buse JB, Nielsen LL, et al. Exenatide effects on diabetes, obesity, cardiovascular risk factors and hepatic biomarkers in patients with type 2 diabetes treated for at least 3 years. *Curr Med Res Opin* 2008; 24: 275–286.
102. Nauck M, Frid A, Hermansen K, et al. Efficacy and safety comparison of liraglutide, glimepiride, and placebo, all in combination with metformin, in type 2 diabetes: the LEAD (Liraglutide Effect and Action in Diabetes)-2 study. *Diabetes Care* 2009; 32: 84–90.
103. Marre M, Shaw J, Brändle M, et al.; LEAD-1 SU study group. Liraglutide, a once-daily human GLP-1 analogue, added to a sulphonylurea over 26 weeks produces greater improvements in glycaemic and weight control compared with adding rosiglitazone or placebo in subjects with type 2 diabetes. *Diabet Med* 2009; 26: 268–278.
104. Seino Y, Rasmussen MF, Zdravkovic M, Kaku K. Dose-dependent improvement in glycemia with once-daily liraglutide without hypoglycemia or weight gain: a double-blind, randomized controlled trial in Japanese patients with type 2 diabetes. *Diabetes Res Clin Pract* 2008; 81: 161–168.
105. Garber A, Henry R, Ratner R, et al.; LEAD-3 (Mono) Study Group. Liraglutide versus glimiperide monotherapy for type 2 diabetes (LEAD-3 Mono): a randomised, 52-week, phase III, double-blind, parallel-treatment trial. *Lancet* 2009; 373: 473–481.
106. Blonde L, Russell-Jones D. The safety and efficacy of liraglutide with or without oral antidiabetic drug therapy in type 2 diabetes: an overview of the LEAD 1-5 studies. *Diabetes Obes Metab* 2009; 11(suppl 3): 26–34.
107. Drucker DJ, Nauck MA. The incretin system: glucagon-like peptide-1 receptor agonists and dipeptidyl peptidase-4 inhibitors in type 2 diabetes. *Lancet* 2006; 368: 1696–1705.
108. Ayoub WA, Kumar AA, Naguib HS, Taylor HC. Exenatide-induced acute pancreatitis. *Endocr Pract* 2010; 16: 80–83.
109. Drucker JD, Sherman SI, Gorelick FS, et al. Incretin-based therapies for the treatment of type 2 diabetes: evaluation of the risks and benefits. *Diabetes Care* 2010; 33: 428–433.
110. Nauck MA, Meininger G, Sheng D, et al. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor, sitagliptin, compared with the sulfonylurea, glipizide, in patients with type 2 diabetes inadequately controlled on metformin alone: a randomized, double-blind, non-inferiority trial. *Diabetes Obes Metab* 2007; 9: 194–205.

111. Raz I, Hanefeld M, Xu L, et al. Efficacy and safety of the dipeptidyl peptidase-4 inhibitor sitagliptin as monotherapy in patients with type 2 diabetes mellitus. *Diabetologia* 2006; 49: 2564–2571.
112. Goldstein B, Feinglos M, Lunceford JK, et al. Effect of initial combination therapy with sitagliptin, a dipeptidyl peptidase-4 inhibitor, and metformin on glycemic control in patients with type 2 diabetes. *Diabetes Care* 2007; 30: 1979–1987.
113. Richter B, Bandeira-Echtler E, Bergerhoff K, Lerch C. Emerging role of dipeptidyl peptidase-4 inhibitors in the management of type 2 diabetes. *Vasc Health Risk Manag* 2008; 4: 753–768.
114. Amori RE, Lau J, Pittas AG. Efficacy and safety of incretin therapy in type 2 diabetes: systematic review and meta-analysis. *JAMA* 2007; 298: 194–206.
115. Young A. Tissue expression and secretion of amylin. *Adv Pharmacol* 2005; 52: 19–45.
116. Juhl CB, Porksen N, Sturis J, et al. High-frequency oscillations in circulating amylin concentrations in healthy humans. *Am J Physiol Endocrinol Metab* 2000; 278: E484–E490.
117. Fineman MS, Koda JE, Shen LZ, et al. The human amylin analog, pramlintide, corrects postprandial hyperglucagonemia in patients with type 1 diabetes. *Metabolism* 2002; 51: 636–641.
118. Kong MF, Stubbs TA, King P, et al. The effect of single doses of pramlintide on gastric emptying of two meals in men with IDDM. *Diabetologia* 1998; 41: 577–583.
119. Chapman I, Parker B, Doran S, et al. Effect of pramlintide on satiety and food intake in obese subjects and subjects with type 2 diabetes. *Diabetologia* 2005; 48: 838–848.
120. Ratner RE, Want LL, Fineman MS, et al. Adjunctive therapy with the amylin analogue pramlintide leads to a combined improvement in glycemic and weight control in insulin-treated subjects with type 2 diabetes. *Diabetes Technol Ther* 2002; 4: 51–61.
121. Hollander PA, Levy P, Fineman MS, et al. Pramlintide as an adjunct to insulin therapy improves long-term glycemic and weight control in patients with type 2 diabetes: a 1-year randomized controlled trial. *Diabetes Care* 2003; 26: 784–790.
122. Riddle M, Frias J, Zhang B, et al. Pramlintide improved glycemic control and reduced weight in patients with type 2 diabetes using basal insulin. *Diabetes Care* 2007; 30: 2794–2799.

MOUNT SINAI JOURNAL OF MEDICINE 77:549–558, 2010

SPECIAL FEATURE

Usefulness of Anti-Infective Lock Solutions for Catheter-Related Bloodstream Infections

Esther Y. Kim, PharmD, Patricia Saunders, PharmD, and Nina Yousefzadeh, PharmD

The Mount Sinai Medical Center, New York, NY

OUTLINE

BACKGROUND
ANTIBIOTIC LOCK SOLUTIONS
INFECTIOUS DISEASES SOCIETY OF AMERICA GUIDELINES: INDICATIONS FOR ANTIBIOTIC LOCK THERAPY
VANCOMYCIN CATHETER-RELATED ANTIBIOTIC LOCK TREATMENT
GENTAMICIN CATHETER-RELATED ANTIBIOTIC LOCK TREATMENT
CEFAZOLIN CATHETER-RELATED ANTIBIOTIC LOCK TREATMENT
NONANTIMICROBIAL CATHETER-RELATED LOCK TREATMENT
MOUNT SINAI HOSPITAL PROTOCOL
CONCLUSION

ABSTRACT

Central venous catheters are commonly used for the administration of medications in both the inpatient and outpatient settings, but their use carries the risk of developing catheter-related bloodstream infections, increasing morbidity and mortality. Guidelines from the Infectious Diseases Society of America include use of antibiotic lock solutions as a therapeutic option for intraluminal infections when the device is difficult to remove, but these lock solutions are not indicated routinely for the prophylaxis of catheter-related bloodstream infections. Because catheter removal is not always possible in patients who require chronic catheterization or who have multiple catheters, antibiotic lock therapy can be utilized with the goal of sparing removal of the catheter while sterilizing the lumen. The use of an antibiotic lock solution is an appealing option for localized infections because it is easy to use, is associated with a low cost, and is usually without systemic side effects if the catheter remains closed. Patients with long-term or short-term catheters with no signs of exit-site or tunnel infection for whom salvage is the goal may benefit from this technique when treated with systemic antibiotics. Consider antibiotic lock therapy only in uncomplicated infections, typically nontunneled catheter or port infections with coagulase-negative staphylococci or gram-negative bacilli. Generally, a short-term regimen of 7 to 14 days is recommended. Several studies have supported antimicrobial lock therapy with vancomycin, gentamicin, cefazolin, or ethanol. Antibiotic locks have the potential to positively impact mortality and cost savings. A protocol at the Mount Sinai Medical Center is in development. ***Mt Sinai J Med 77:549–558, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: antibiotic lock therapy, bacteremia, catheter-related infections.

Address Correspondence to:

Esther Y. Kim
Mount Sinai Medical Center
New York, NY
Email: esther.kim@mountsinai.org

Catheter-related bloodstream infections (CRBSIs) are the leading contributor to healthcare-associated infections. Approximately 250,000 CRBSIs occur yearly in the United States out of the 150 million intravascular devices used.[1,2] This correlates with an increased length of stay in the intensive care unit by approximately 2.5 days and increased inpatient stay by 7.5 days.[1] Risk factors for developing CRBSIs include duration of central venous catheter (CVC), coexisting illnesses, patient location, hospital size,

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20213

© 2010 Mount Sinai School of Medicine

number of lumens, catheter type, use for parenteral nutrition, emergent placement, and number of manipulations per day. The cost per infection is estimated to be \$34,508 to \$56,000, and the annual cost of caring for patients with CVC-associated bloodstream infections ranges from \$296 million to \$2.3 billion.[3] Thus, the cost of CVC-associated bloodstream infections is substantial in terms of morbidity and expended financial resources.

CRBSIs are difficult to treat because they are caused by microorganisms that can attach to catheter surfaces by creating biofilms. Biofilms render antibiotics less effective because it is difficult for antibiotics to penetrate the polysaccharide layer, and biofilms are unaffected by phagocytosis. The recommended treatment for an infected CVC is line removal, but when it cannot be removed, instilling an antibiotic at a concentration of 10× – 1000× higher than the minimum inhibitory concentration (MIC), otherwise known as antibiotic lock therapy (ALT), may be used for catheter salvage.

Guidelines from the Infectious Diseases Society of America (IDSA) include use of antibiotic lock solutions as a therapeutic option for intraluminal infections when the device is difficult to remove, but these lock solutions are not indicated routinely for the prophylaxis of CRBSIs. Catheter removal is necessary for catheter tunnel or implanted port pocket infection, septic shock or complicated infection, positive blood cultures for fungi or drug-resistant bacteria, or reinfection with the same microorganism after previous treatment. Managing these infections successfully requires appropriately assessing whether the catheter should be removed or preserved using ALT, as well as selecting an antibiotic regimen based on the bacteria and its susceptibility pattern.[4]

> *Managing catheter-related bloodstream infections successfully requires appropriately assessing whether the catheter should be removed or preserved using antibiotic lock therapy, as well as selecting an antibiotic regimen based on the bacteria and its susceptibility pattern.*

The Joint Commission on Accreditation of Healthcare Organizations requires hospitals to implement evidence-based best practices to prevent bloodstream infections caused by both short-term and long-term CVCs, as well as peripherally inserted catheter lines, and provides guidelines for hospitals to meet National Patient Safety Goals.[5] The elements of performance include:

1. Educate staff and practitioners about CRBSI and the importance of their prevention.
2. Prior to catheter insertion, educate patients and families about CRBSI prevention.
3. Implement policies and practices aimed at reducing the risk of CRBSI that meet regulatory requirements, such as those from the Centers for Disease Control and Prevention.
4. Conduct periodic, scheduled risk assessments; monitor compliance; and evaluate the effectiveness of prevention efforts.
5. Use a catheter checklist and standardized protocol for insertion of catheters.
6. Perform hand hygiene prior to insertion or manipulation of catheters.
7. Sites other than the femoral vein should be inserted in adult patients, unless other sites are unavailable.
8. Use a standardized supply cart or kit that contains all components necessary for catheter insertions.
9. Unless contraindicated, use a chlorhexidine-based antiseptic for skin preparation during catheter insertion in patients aged ≥2 months.
10. Use a standardized protocol to disinfect catheter hubs and injection ports prior to accessing the ports.
11. Evaluate all catheters routinely and remove nonessential catheters.

Recent evidence has shown that CRBSIs are preventable, resulting in CRBSI now being considered a preventable complication by the Centers for Medicare and Medicaid Services.[1] Beginning in October 2008, it no longer reimburses hospitals for expenses related to catheter-related infections. Without direct reimbursement for CVC infection prevention, hospitals must fund their own infection-control interventions. Due to limited data and hospitals' limited monetary resources, the outcome of these prevention programs may not be able to provide the greatest societal benefit. Although the IDSA recommends against the routine use of prophylactic antibiotic lock solutions, several studies have described efficacy in using lock solutions as an option to reduce healthcare costs and the incidence of catheter-related infections.[6]

> *Although the Infections Diseases Society of America recommends against the routine use of prophylactic antibiotic lock solutions, several studies have described efficacy in using lock solutions as an option to reduce healthcare costs and the incidence of catheter-related infections.*

BACKGROUND

For microorganisms to cause a CRBSI, the bacteria or fungi must initially gain access to the extraluminal or intraluminal surface of the device. Microorganisms gain access to an implanted device via the following mechanisms[7]: skin organisms invade the percutaneous tract during or at the time of catheter insertion, microorganisms contaminate the catheter hub when inserted or manipulated (medication infusions, tubing or cap changes, flushing or blood sampling), or organisms are carried hematogenously to the implanted device from a remote source of a local infection (ie, pneumonia).

Once they gain access, they become incorporated into a biofilm, a self-protective adhesive matrix that allows sustained colonization and ultimately hematogenous growth. Biofilms are produced by a structured community of bacterial cells in a hostile environment as a means of protection. Biofilm formation within the catheter lumen limits penetration of antimicrobial solutions, and standard intravenous (IV) antibiotic therapy may not reach concentrations high enough to reduce the bacterial burden. Bloodstream infections are caused by the detachment of cells or clump of cells from the biofilm, the production of endotoxins, or ineffectiveness of the host to fight the infection. Instillation of an antibiotic lock may eradicate biofilms in vitro.

ANTIBIOTIC LOCK SOLUTIONS

The rationale behind antibiotic lock therapy is to salvage the catheter if possible, because catheters are difficult to implant and remove, and are expensive. Furthermore, vascular sites are limited, especially in the pediatric population. Because catheter removal is not always possible in patients who require chronic catheterization or who have multiple catheters, ALT can be utilized with the goal of sparing removal of the catheter while sterilizing the lumen. This technique has been successfully employed in patients with cancer and acquired immunodeficiency syndrome, and in those requiring hemodialysis or total parenteral nutrition.[8] The goals of therapy are to eradicate the organism from the infected line, prevent colonization of a sterile line, salvage the catheter, and reduce morbidity and the costs of managing these infections.

> *Because catheter removal is not always possible in patients who require chronic catheterization or who have multiple catheters, antibiotic lock therapy can be utilized with the goal of sparing removal of the catheter while sterilizing the lumen.*

The model for ALT was developed in the late 1980s as a derivation of the heparin lock solutions. The idea hypothesized that high peak concentrations (10× – 1000× higher than used systemically) of bactericidal antibiotic solutions could effectively sterilize a catheter and therefore decrease catheter-related sepsis.[3] This model entails a highly concentrated antibiotic solution instilled into the catheter lumen and allowing the solution to reside in the lumen for several hours between infusions. This technique is similar to heparin lock solutions, but with the intention of lumen sterilization.

The antibiotic choice depends on the pharmacodynamic properties of the antibiotic, the suspected pathogen infecting the catheter lumen, and the specific characteristics of the bacteria likely to be causing an infection. The selection of a bactericidal over bacteriostatic agent seems appropriate and may be guided by the susceptibility testing results. Antibiotics such as aminoglycosides and fluoroquinolones exhibit optimal bacterial killing at high concentrations. As a result of the postantibiotic effect of these drugs, it may not be necessary to maintain high concentrations throughout a dosing interval or dwelling time within a catheter. In comparison, beta-lactams and vancomycin exhibit time-dependent killing, which requires a maintenance concentration above the targeted organism's MIC for the majority of the dosing interval. Lock solutions are most effective when high concentrations are maintained for an extended time period in order to maximize bactericidal activity and biofilm penetration. The IDSA guidelines recommend that the dwelling time of the solution should not exceed 48 hours before reinstallation of the lock solution.[4] For ambulatory patients with femoral catheters, the guidelines recommend reinstallation of the lock solution at 24 hours.[4] There are studies that suggest antibiotic-heparin lock solutions with dwell times of up to 10 days are feasible, but the guidelines recommend that for hemodialysis patients, the lock solution can be replaced after every dialysis session and should not exceed a 48-hour dwelling time.[4,9]

ALT with or without heparin has been studied. In antibiotic-anticoagulation lock combinations, heparin prevents adhesion of proteins such as fibrin, fibronectin, and fibrinogen to the catheter, to which organisms such as *Staphylococcus aureus* and candida species like to adhere. The combination solutions contain the desired antimicrobial concentration and are usually mixed with a range of 50–5000 units of heparin in sufficient volume with normal saline, dextrose, or lactated Ringer's to fill the catheter lumen.[2] Concerns with heparin include issues with drug stability, large concentrations inadvertently causing systemic heparinization, and introducing an anticoagulant to actively hypocoagulable patients.

INFECTIOUS DISEASES SOCIETY OF AMERICA GUIDELINES: INDICATIONS FOR ANTIBIOTIC LOCK THERAPY

In July 2009, the IDSA published new guidelines for the management of intravascular catheter-related infections. The most widespread organisms causing CRBSI in patients with percutaneously inserted noncuffed catheters, in order of highest to lowest prevalence, are coagulase-negative staphylococci (ie, *Staphylococcus epidermidis*), *S. aureus*, *Candida* species, and enteric gram-negative bacilli.[4] For patients with surgically implanted catheters and peripherally inserted CVCs, the most widespread organisms causing infections include coagulase-negative staphylococci (CoNS), enteric gram-negative bacilli, *S. aureus*, and *Pseudomonas aeruginosa*.[4]

ALT is indicated for both long-term CVC and short-term or arterial catheter infections. For patients with uncomplicated CRBSIs caused by CoNS or enteric gram-negative bacilli that involve long-term catheters, without any signs of exit-site or tunnel infection and in whom catheter salvage is the goal, ALT can be used.[4] The guidelines recommend treatment with systemic antibiotics for 10 to 14 days plus ALT for 10 to 14 days for CoNS or 7 to 14 days for gram-negative bacilli.[4] For short-term CVC or arterial catheters that are retained, systemic antibiotics for 5 to 7 days and ALT for 10 to 14 days is indicated for positive CoNS cultures.[4] Success rates of catheter preservation in hemodialysis patients diagnosed with CRBSI with ALT are 87% to 100% for gram-negative pathogens, 75% to 84% for *S. epidermidis*, and 40% to 55% for *S. aureus* infections.[4] Therefore, the guidelines recommend against the use of ALT for complicated infections or infections caused by *S. aureus* or *Candida* species (Table 1 and Table 2).

The use of an antibiotic lock solution is an appealing option for localized infections because it is easy to use, is associated with a low cost, and is usually without systemic side effects if the catheter remains closed. Potential disadvantages may be the loss of antibiotic activity with dwelling time and possible treatment delay if the antibiotic lock fails.[8] Efficacy of the antibiotic lock solutions may diminish over time because antibiotic concentrations rapidly decrease in the catheter lumen. The solution warrants a specific antimicrobial minimum inhibitory concentration during the entire dwell time.[4]

> *The use of an antibiotic lock solution is an appealing option for localized infections because it is easy to use, is associated with a low cost, and is usually without systemic side effects if the catheter remains closed.*

VANCOMYCIN CATHETER-RELATED ANTIBIOTIC LOCK TREATMENT

Several studies have investigated the use of antibiotic lock solutions. Bacteremia is a frequent complication among hemodialysis patients using tunneled dialysis catheters. Kidney Disease Outcomes Quality Initiative guidelines recommend treatment of CRBSI with systemic antibiotics plus catheter replacement. In one study, patients with suspected dialysis catheter–associated bacteremia received loading doses of vancomycin 20 mg/kg and ceftazidime 1 g, followed by maintenance doses of vancomycin 500 mg and ceftazidime 1 g during each consecutive dialysis session. Among the 83 episodes of CRBSI, 15

Table 1. *Approach to the Management of Patients with Short-Term Central Venous Catheter–Related or Arterial Catheter–Related Bloodstream Infection*[5].

Diagnosis	Treatment	ALT Indicated?
Complicated infections		
Suppurative thrombophlebitis, endocarditis, or osteomyelitis	Remove catheter and treat with systemic antibiotics for 4 to 6 weeks; 6 to 8 weeks for osteomyelitis in adults	No
Uncomplicated infections		
CoNS	Remove catheter and treat with systemic antibiotics for 5 to 7 days; If catheter is retained: treat with systemic antibiotics and antibiotic lock therapy for 10 to 14 days	Yes; use salvage therapy
S. aureus	Remove catheter and treat with systemic antibiotics for at least 14 days	No
GNR	Remove catheter and treat with systemic antibiotics for 7 to 14 days	No
Candida species	Remove catheter and treat with systemic antibiotics for 14 days after the first negative blood culture	No

Abbreviations: ALT, antibiotic lock therapy; CoNS, coagulase-negative *Staphylococcus*; GNR, gram-negative rods; *S. aureus*, *Staphylococcus aureus*.

were excluded from the protocol because the blood cultures grew *Enterococcus* species, and the patients were referred for a routine catheter exchange. Forty-seven episodes of catheter-related bacteremia in hemodialysis patients were treated with 3 weeks of systemic antibiotics (vancomycin + ceftazidime) plus the ALT protocol of vancomycin 2.5 mg/mL + heparin 2500 units/mL. The ALT solution combination

Table 2. *Approach to the Treatment of a Patient with a Long-Term Central Venous Catheter or a Port-Related Bloodstream Infection*[5].

Diagnosis	Treatment	ALT Indicated?
Complicated infections		
Tunnel infection, port abscess	Remove CVC or port and treat with systemic antibiotics for 7 to 14 days	No
Septic thrombosis, endocarditis, or osteomyelitis	Remove CVC or port and treat with systemic antibiotics for 4 to 6 weeks; 6 to 8 weeks for osteomyelitis in adults	No
Uncomplicated infections		
CoNS	Treat with systemic antibiotics and antibiotic lock therapy for 10 to 14 days Remove if clinical deterioration or relapse of bacteremia	Yes; use salvage therapy
S. aureus	Remove CVC or port and treat with systemic antibiotics for 4 to 6 weeks; 10 to 14 days if the patient does NOT have: diabetes mellitus, immunosuppression, prosthetic valve, endocarditis, suppurative thrombophlebitis on TEE, fever or bacteremia for <12 hours	No
GNR	Treat with systemic antibiotics and antibiotic lock therapy for 7 to 14 days Remove if clinical deterioration or relapse of bacteremia	Yes; use salvage therapy
Candida species	Remove catheter and treat with systemic antibiotics for 14 days after the first negative blood culture	No

Abbreviations: ALT, antibiotic lock therapy; CVC, central venous catheter; CoNS, coagulase-negative *Staphylococcus*; GNR, gram-negative rods; *S. aureus*, *Staphylococcus aureus*; TEE, transesophageal echocardiogram.

was successful (patients remained asymptomatic with negative cultures and the catheter was salvaged) in 33 cases (70%) but failed in 14 cases (30%).[10] *S. epidermidis* and *S. aureus* (79% methicillin-resistant *S. aureus*) were the primary organisms seen, and 87% and 40%, respectively, had positive cultures. The median catheter survival time without infection following the initial episode of CRBSI was about twice as long in the ALT protocol group as compared with the routine catheter replacement group (154 versus 71 days, $P = 0.02$).[10] The success rate varied distinctly according to the type of infection and the causative organism, being quite high for *S. epidermidis* or gram-negative rods, but low for *S. aureus*. The vancomycin-ceftazidime ALT protocol was successful in about two-thirds of episodes. The findings suggest that the overall success rate in an ALT protocol for hemodialysis patients will depend on the frequency of causative bacterial pathogens.[10]

A randomized, double blind, placebo-controlled study used ALT in the treatment of CRBSI in patients with long-term intravascular devices. All patients received systemic antimicrobial therapy through the catheter according to the IDSA guidelines. In addition, patients received either ALT or placebo for 7 to 14 days. The placebo syringes were filled with heparin 100 units/mL, and the ALT syringes with a solution of heparin and ceftazidime (for gram-negative bacteria) or vancomycin (for gram-positive bacteria), both at a concentration of 500 mg/mL. The primary endpoint was failure to cure CRBSI, which was defined as removal of the catheter for any reason during treatment, relapse of bacteremia with the same microorganism, or death during the initial ALT phase or due to a catheter infection at any time during the 6-month follow-up. Forty-four patients were included in the intention-to-treat analysis, with a median duration of antibiotic lock instillation of 20 hours per day for a total of 11 days. On study day 180, failure to cure the CRBSI or relapse with the same strain occurred in 20 of 44 patients (45.5%), 33% (7 of 21) in the ALT arm and 57% (13 of 23) in the placebo arm (hazard ratio 0.55, $P = 0.10$).[11] Relapse with the same strain occurred in 3 of 21 (14.3%) patients in the ALT group versus 9 of 23 (39.1%) in the placebo group ($P = 0.06$).[11] Although not statistically significant, the study showed that in patients with a long-term intravascular device–related bloodstream infection, the use of an antibiotic lock solution, in addition to IV antibiotics, reduced the failure to cure CRBSI from 58% to 33%. ALT prevented relapses of CRBSI with the same strain, but did not increase the cure rate of the first-episode CRBSI. The study further supports the recommendations in the IDSA guidelines for ALT for the treatment of CRBSI.

The use of ALT for CRBSI associated with long-term intravascular devices used for chemotherapy or parenteral nutrition was investigated during a 44-month period. In this randomized controlled trial, 801 long-term intravascular devices were placed in 105 patients, and 127 episodes of CRBSI were documented. Forty-eight episodes were managed with IV antibiotics; of these, 19 episodes were treated with ALT and 29 episodes were treated with systemic antibiotics. The median time of systemic antibiotic administration was 11 days. In addition to the systemic antibiotics, an ALT combination solution of vancomycin 2 mg/mL (for gram-positive microorganisms) + heparin 20 units/mL (and ciprofloxacin 2 mg/mL or gentamicin 2 mg/mL for gram-negative microorganisms) was administered. The primary endpoint was failure to cure CRBSI. Treatment success was observed in 84% of patients in the ALT combination group versus 65% in the control group ($P = 0.06$).[12] Specific analysis of CoNS episodes showed that the combination cured 93% of patients versus 79% of patients treated with the control ($P = 0.36$).[12] ALT was successful in 84% of all episodes and in 94% caused by CoNS. Treatment failed in two-thirds of patients with *S. aureus*–associated CRBSI. The study concludes that ALT is effective in treating CRBSI caused by CoNS in patients with long-term intravascular devices (Table 3).

GENTAMICIN CATHETER-RELATED ANTIBIOTIC LOCK TREATMENT

Several studies looked at the role of gentamicin-heparin lock solutions. Fifty patients receiving hemodialysis with a confirmed CRBSI were randomized to receive standard heparin 5000 units/mL or heparin + gentamicin 5 mg/mL. The CRBSI was treated with an empiric IV regimen of vancomycin and gentamicin. The mean time to first episode of CRBSI was significantly longer in the gentamicin lock–treated group (130 days) compared with heparin lock alone (85.3 days) ($P = 0.03$).[13] The most common microorganisms cultured were *S. epidermidis, S. aureus,* and gram-negative bacilli. For catheter infection rates, 1 episode in the ALT group versus 10 episodes in the placebo, consistent with 0.3 episodes per 1000 catheter days and 4 per 1,000 catheter days, significantly favored the ALT group ($P = 0.03$).[13] All random gentamicin levels were <0.2 mg/L, and adverse events were not reported or seen in either group. The study demonstrated

Table 3. *Summary of Efficacy Endpoints of Vancomycin Antibiotic Lock Studies.*

Study Population	Treatment	Catheter	Endpoints and Results
Forty-seven episodes of catheter-related bacteremia in hemodialysis patients[8]	Episodes were treated with vancomycin 20 mg/kg and ceftazidime 1 g during dialysis session for 3 weeks	Vancomycin 2.5 mg/mL + heparin 2500 units/mL versus routine catheter replacement	Catheter survival time without infection following the initial infection episode with ALT was longer at 154 days versus routine catheter replacement at 71 days; $P = 0.02$
Randomized placebo-controlled trial of 44 CRBSI patients with long-term intravascular devices: CVC or implantable ports[9]	Patients were treated with systemic antibiotic therapy through the catheter according to the IDSA guidelines	Vancomycin or ceftazidime, both at 500 mg/mL, versus heparin 100 units/mL	ALT reduced the failure to cure CRBSI from 57% to 33%; $P = 0.10$ Relapse with the same strain was reduced from 39.4% to 14.3%; $P = 0.06$
Twenty-nine CRBSI episodes in patients with long-term intravascular devices: Hickman or implantable reservoirs[10]	Episodes were treated with systemic antibiotics for a median time of 11 days	Vancomycin 2 mg/mL (gram-positive organisms) + heparin 20 units/mL versus heparin	Treatment success was observed in 84% in the ALT group versus 65% in the control group; $P = 0.06$

Abbreviations: ALT, antibiotic lock therapy; CRSBI, catheter-related bloodstream infection; CVC, central venous catheter; IDSA, Infectious Diseases Society of America.

a markedly reduced incidence of CRBSI without adverse events, and an added benefit appeared on treatment costs due to a reduction in epoetin requirements in hemodialysis patients using tunneled hemodialysis catheters.

Sixty-two patients on hemodialysis were treated with 3 weeks of IV antibiotics and ALT of gentamicin 1 mg/mL + heparin 2500 units/mL. Approximately 20% of hemodialysis patients in the United States use a tunneled central vein catheter for vascular access, and bacteremia occurs at a frequency of 2 to 5.5 per 1000 catheter days.[14] The efficacy of an ALT protocol in conjunction with systemic antibiotics was studied in patients with dialysis catheter–associated bacteremia without catheter removal. Treatment success was defined as the resolution of fever and negative cultures for 1 week after completion of the protocol. Of the 62 episodes where the catheter was salvaged, 40 cases (64.5%) were protocol successes where the patient remained asymptomatic with negative cultures and the catheter was salvaged; the remaining 22 cases (35.5%) were considered protocol failures.[14] No statistical difference was seen in infection-free survival or overall survival compared with patients whose infections were treated by catheter replacement.[14] One case of methicillin-resistant *S. aureus* was unsuccessfully treated; however, most of the organisms were staphylococcus species and gram-negative bacilli. ALT in dialysis CRBSI can be eradicated in two-thirds of cases without requiring removal of the infected catheter.[14] However, the probability of a subsequent infection was similar in both treatment groups. The study suggests that the use of ALT, in addition to IV antibiotics, offers an advantage over routine catheter exchange (Table 4).

CEFAZOLIN CATHETER-RELATED ANTIBIOTIC LOCK TREATMENT

The majority of the cefazolin ALT studies were in vitro, demonstrating stability and efficacy; however, one study observed the incidence of catheter infections in patients with temporary catheters. ALT was instituted using a combination cefazolin-heparin lock solution in 17 patients with 25 catheters. Patients were treated with a lock solution of cefazolin 5 mg/dL and heparin 2500 units/mL. Mupirocin was topically applied to the peripheral area of the catheter exit site. The control group consisted of 9 patients with 15 catheters who had catheter replacement instead. Treatment endpoints were CRBSI, exit-site infections, or patient death. The catheter infection rate was reduced from 12.7 times per 1000 catheter days to 5.02 times per 1000 catheter days in patients with jugular vein catheters.[15] The total catheter-related infection rate was 14.9 times per 1000 catheter days in the control group and 4.1 times per 1000 catheter days in the study group.[15] The use of ALT and topical antibiotics reduced the incidence of temporary

Table 4. *Summary of Efficacy Endpoints of Gentamicin Antibiotic Lock Studies.*

Study Population	Treatment	Catheter	Endpoints and Results
Fifty patients receiving hemodialysis with a confirmed CRBSI[11]	Systemic antibiotic regimen of vancomycin and gentamicin	Gentamicin 5 mg/mL + heparin 5000 units/mL versus heparin 5000 units/mL	Mean time to first episode of CRBSI, 130 days versus 85.3 days; $P = 0.03$
Sixty-two patients with dialysis-associated bacteremia[12]	Systemic antibiotic for 3 weeks	Gentamicin 1 mg/mL + heparin 2500 units/mL versus catheter replacement	Resolution of fever and negative cultures for 1 week post protocol, 64.5% versus 35.5%

Abbreviation: CRSBI, catheter-related bloodstream infection.

catheter-related infections, especially in those with femoral catheters.

NONANTIMICROBIAL CATHETER-RELATED LOCK TREATMENT

Similar to ALT, an ethanol lock fills and closes the catheter lumen with an ethanol solution. Because ethanol acts by a nonspecific protein denaturation, the advantages of using ethanol over antibiotics are eradicating bacterial colonization irrespective of the sensitivity of the organism and a decreased risk of developing antimicrobial resistance. Ethanol is not compatible with heparin, so it should not be combined with heparin flushes during therapy. A 50% or 70% ethanol lock concentration is adequate for prevention of infection because a concentration greater than 40% is able to inhibit bacterial growth in established biofilms.[7] However, a 70% ethanol lock must dwell in the catheter for at least 2 hours to be effective.[7] Heparin should not be administered concomitantly with ethanol lock therapy due to formation of a precipitate.

Heparin should not be administered concomitantly with ethanol lock therapy due to formation of a precipitate.

The effectiveness of antibiotics and ethanol as catheter lock techniques has been compared to treat catheter colonization. The effectiveness of catheter lock solutions was assessed by determining the minimal biofilm-eradication concentration of antimicrobial agents (oxacillin, gentamicin, vancomycin, ciprofloxacin, and rifampin) and ethanol. Ethanol was superior to all of the conventional antibiotics, and a time-kill study and confocal laser scanning microscopy revealed that exposure to 40% ethanol for 1 hour was sufficient to kill CoNS biofilm cells.[16] The study suggests that there is good evidence to support the use of low-concentration ethanol for short-term use, as opposed to high ALT concentrations, to treat CRBSIs. However, potential for clinically important systemic side effects due to the release of plastic components after ethanol lock use needs to be evaluated and compared with the leaching of chemicals associated with other lock solutions.[17]

When ethanol lock solutions are considered, the effect of ethanol on the mechanical and structural integrity of the catheter must also be considered. The effect of prolonged exposure to 70% ethanol on the mechanical integrity of 2 types catheters commonly used for long-term vascular access was studied. For 10 weeks, 1 set of catheters (15 polyetherurethane and 16 silicone catheters) was locked with 70% ethanol lock solution, whereas the control catheters were left empty without any solution. The catheters in both groups were immersed in a salt solution at a temperature of 37°C, to simulate the effect of human bloodstream. A 70% ethanol lock solution had minimal impact on the mechanical properties (force-at-break, stress, strain, modulus of toughness) and marginal reduction in elasticity compared with the control group.[18] The findings support the safety of using ethanol lock solutions with selected polyetherurethane and silicone catheters in clinical practice.

Chelating agents such as ethylene diamine tetraacetic acid or citrate locks may reduce catheter infections and may have an advantage because they will not generate antimicrobial resistance. They show anticoagulant activity (similar to heparin) and show enhanced activity against organisms embedded in biofilm.[19] Chelating agents act on calcium and magnesium, and the concern is their risk of cardiac arrhythmias if there is leakage and entry into the systemic circulation.[2]

MOUNT SINAI HOSPITAL PROTOCOL

A protocol for the use of antibiotic and ethanol locks is in development and is being adapted from the 2009 IDSA guidelines. Because the most effective method to eliminate an infection is removal of the catheter, catheter salvage should be avoided and the line simply removed in the following situations: short-term CRBSI, exit-site or tunnel infection, severe sepsis or shock, persistent bacteremia despite 72 hours of appropriate antibacterial therapy, evidence of metastatic infections such as endocarditis or osteomyelitis, presence of intravascular hardware, *S. aureus, P. aeruginosa*, fungi, mycobacteria, miscellaneous pathogens that are difficult to eradicate (eg, *Bacillus, Micrococcus*), or relapses of any bacteremia.

In cases when ALT is feasible, management guidelines for catheter salvage recommend the continuation of IV antibiotics until the clinical signs and symptoms of the infection have resolved and blood cultures are negative. ALT should then be implemented for 14 days from the date of negative blood cultures. If the ALT is not feasible, then IV antibiotics through the colonized catheter should be continued for 14 days.

The dwell time for ALT solutions should be at least 6 hours twice per day (total of at least 12 hours). It is important to aspirate the antibiotic solution prior to using the line for medication administration. ALT is considered a failure if cultures are still positive after 48 hours of ALT, or other indications for catheter removal develop at any time.

CONCLUSION

ALT in conjunction with systemic antibiotics can treat CRBSIs while sparing the catheter. Patients with long-term or short-term catheters with no signs of exit-site or tunnel infection and for whom salvage is the goal may benefit from this technique when treated with systemic antibiotics. Consider ALT only in uncomplicated infections, typically nontunneled catheter or port infections with CoNS or gram-negative bacilli. Generally, a short-term regimen of 7 to 14 days is recommended. Several studies have supported ALT with vancomycin, gentamicin, cefazolin, or ethanol. ALT has the potential to positively impact mortality and cost savings. A protocol at the Mount Sinai Medical Center is in development.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Goede MR, Coopersmith CM. Catheter-related bloodstream infections. *Surg Clin N Am* 2009; 89: 463–474.
2. Del Pozo JL. Role of antibiotic lock therapy for the treatment of catheter-related bloodstream infections. *Int J Artif Organs* 2009; 32: 678–688.
3. Bestul MB, Vandenbussche HL. Antibiotic lock technique: review of the literature. *Pharmacotherapy* 2005; 25: 211–227.
4. Mermel LA, Allon M, Bouza E, et al. Clinical practice guidelines for the diagnosis and management of intravascular catheter-related infection: 2009 update by the Infectious Diseases Society of America [published corrections appear in *Clin Infect Dis* 2010; 50: 457 and 2010; 50: 1079]. *Clin Infect Dis* 2009; 49: 1–45.
5. The Joint Commission on Accreditation of Healthcare Organizations. Accreditation Program: Hospital. National Patient Safety Goals, 2009.
6. Perencevich EN, Pittet D. Preventing catheter-related bloodstream infections: thinking outside the checklist. *JAMA* 2009; 301: 1285–1287.
7. Bell AL, Gu X, Burczynski FJ, Vercaigne LM. Ethanol/trisodium citrate for hemodialysis catheter lock. *Clin Nephrol* 2004; 62: 369–373.
8. Carratalà J. The antibiotic lock technique for therapy for highly needed infected catheters. *Clin Microbiol Infect* 2002; 8: 282–289.
9. Anthony TU, Rubin LG. Stability of antibiotics used for antibioticlock treatment of infections of implantable venous devices (ports). *Antimicrob Agents Chemother* 1999; 43: 2074–2076.
10. Poole CV, Carlton D, Bimbo L, Allon M. Treatment of catheter-related bacteraemia with an antibiotic lock protocol: effect of bacterial pathogen. *Nephrol Dial Transplant* 2004; 19: 1237–1244.
11. Rijnders BJ, Van Wijngaerden E, Vandecasteele SJ, et al. Treatment of long-term intravascular catheter-related bacteremia with antibiotic lock: randomized, placebo-controlled trial. *J Antimicrob Chemother* 2005; 55: 90–94.
12. Fortún J, Grill F, Martín-Dávila P, et al. Treatment of long-term catheter-related bacteremia with antibiotic lock therapy. *J Antimicrob Chemother* 2006; 58: 816–821.
13. McIntyre CW, Hulme LJ, Taal M, Fluck RJ. Locking of tunneled hemodialysis catheters with gentamicin and heparin. *Kidney Int* 2004; 66: 801–805.
14. Krishnasami Z, Carlton D, Bimbo L, et al. Management of hemodialysis catheter-related bacteremia with an adjunctive antibiotic lock solution. *Kidney Int* 2002; 61: 1136–1142.
15. Chiou PF, Chang CC, Wen YK, Yang Y. Antibiotic lock therapy reduces the incidence of temporary catheter-related infections. *Clin Nephrol* 2006; 65: 419–422.
16. Qu Y, Istivan TS, Daley AJ, et al. Comparison of various antimicrobial agents as catheter lock solutions: preference for ethanol in eradication of coagulase-negative staphylococcal biofilm. *J Med Microbiol* 2009; 58: 442–450.
17. Maiefski M, Rupp ME, Hermsen ED. Ethanol lock technique: review of the literature [published correction appears in *Infect Control Hosp Epidemiol* 2010; 31: 202]. *Infect Control Hosp Epidemiol* 2009; 10: 1096–1108.

18. Crinch CJ, Halfmann JA, Crone WC, Maki DG. The effects of prolonged ethanol exposure on the mechanical properties of polyurethane and silicone catheters used for intravascular access. *Infect Control Hosp Epidemiol* 2005; 20: 708–714.

19. Raad II, Hachem RY, Hanna HA, et al. Role of ethylene diamine tetra-acetic acid (EDTA) in catheter lock solutions: EDTA enhances antifungal activity of amphotericin B lipid complex against *Candida* embedded in biofilm. *Int J Antimicrob Agents* 2008; 32: 515–518.

Mount Sinai Journal of Medicine 77:559–569, 2010

SPECIAL FEATURE

Mechanisms of Yogic Practices in Health, Aging, and Disease

Viktoriya Kuntsevich, PhD,[1] William C. Bushell, PhD,[2] and Neil D. Theise, MD[1]

[1]Beth Israel Medical Center of Albert Einstein College of Medicine, New York, NY
[2]Massachusetts Institute of Technology, Anthropology Program, Cambridge, MA; Tibet House US, New York, NY

OUTLINE

Transduction Pathways of Yogic Practices
- **Humoral Signaling**
- **Nervous System Activity**
- **Cell Trafficking**
- **Bioelectromagnetism**

Toward a Changing Paradigm for Investigation

Conclusion

ABSTRACT

Mechanisms underlying the modulating effects of yogic cognitive-behavioral practices (eg, meditation, yoga asanas, pranayama breathing, caloric restriction) on human physiology can be classified into 4 transduction pathways: humoral factors, nervous system activity, cell trafficking, and bioelectromagnetism. Here we give examples of these transduction pathways and how, through them, yogic practices might optimize health, delay aging, and ameliorate chronic illness and stress from disability. We also recognize that most studies of these mechanisms remain embedded in a reductionist paradigm, investigating small numbers of elements of only 1 or 2 pathways. Moreover, often, subjects are not long-term practitioners, but recently trained. The models generated from such data are, in turn, often limited, top-down, without the explanatory power to describe beneficial effects of long-term practice or to provide foundations for comparing one practice to another. More flexible and useful models require a systems-biology approach to gathering and analysis of data. Such a paradigm is needed to fully appreciate the deeper mechanisms underlying the ability of yogic practice to optimize health, delay aging, and speed efficient recovery from injury or disease. In this regard, 3 different, not necessarily competing, hypotheses are presented to guide design of future investigations, namely, that yogic practices may: (1) promote restoration of physiologic setpoints to normal after derangements secondary to disease or injury, (2) promote homeostatic negative feedback loops over nonhomeostatic positive feedback loops in molecular and cellular interactions, and (3) quench abnormal "noise" in cellular and molecular signaling networks arising from environmental or internal stresses. ***Mt Sinai J Med 77:559–569, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: circadian, cytokine, inflammation, integrative medicine, meditation, melatonin, pranayama, vagus nerve, yoga.

> Imagination is more important than knowledge. For while knowledge defines all we currently know and understand, imagination points to all we might yet discover and create. —Albert Einstein

Practitioners of cognitive-behavioral practices (CBPs) of diverse "yogic" traditions, such as yoga asana, meditation, pranayama breathing, and caloric restriction, have made claims regarding their benefits for optimization of health, protection from injury, rejuvenation, and increased life expectancy. It should be kept in mind that although we use the term "yogic," typically indicating derivation from practices historically traced to south Asia, variations of such practices are found in many cultures and traditions around the world, some of which may represent importation

Address Correspondence to:

Viktoriya Kuntsevich, PhD
Nephrology Clinical Research Laboratory
Beth Israel Medical Center
330 East 17th Street,
Silver Bldg, Lab 6S55
New York, NY 10003
Email: vkuntsev@chpnet.org

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI:10.1002/msj.20214

© 2010 Mount Sinai School of Medicine

from that region whereas others are independent developments. However, contemporary research of such practices, which we will discuss in this speculative review, is most often focused on practices from specifically Asian yogic traditions.

TRANSDUCTION PATHWAYS OF YOGIC PRACTICES

Exploration of such physiology-modulating effects of yogic CBPs on health and longevity suggests 4 possible transduction pathways of such global/systemic interventions into site-specific, local effects: humoral factors, nervous system activity, cell trafficking, and bioelectromagnetism (Figure 1).[1] In this article, we will give some recent examples of each of these pathways and point to the often tight interconnections between pathways. However, we also recognize that such examples are simply demonstrations of end effects, but do not really reach to the core of "what yogic practices do" in the process of improving systemic regulation or "restoring balance" to the system. Indeed, most of these investigations, although of great significance, remain in the realm of reductive science. Though an important first step in the exploration of the modulating effects of yogic CBPs, they are incomplete, largely focused on a small number of physiologic events in recently trained practitioners. The leap to a systems-analytic approach is necessary for a comprehensive understanding of the mechanisms whereby yogic CBPs influence, optimize, or restore human physiology.

> *A comprehensive understanding of the mechanisms whereby yogic cognitive-behavioral practices (CBPs) influence, optimize, or restore human physiology requires a leap to a systems-analytic approach. Within this framework, explorations suggest 4 possible transduction pathways of global/systemic yogic interventions into site-specific, local effects: humoral factors, nervous system activity, cell trafficking, and bioelectromagnetism.*

Thus, we will follow these specific examples of each transduction pathway and their interconnectedness, with a presentation of speculative hypotheses that we hope will lead to more detailed and useful explanations of these intriguing and perhaps (literally) vital effects, including optimization of health,

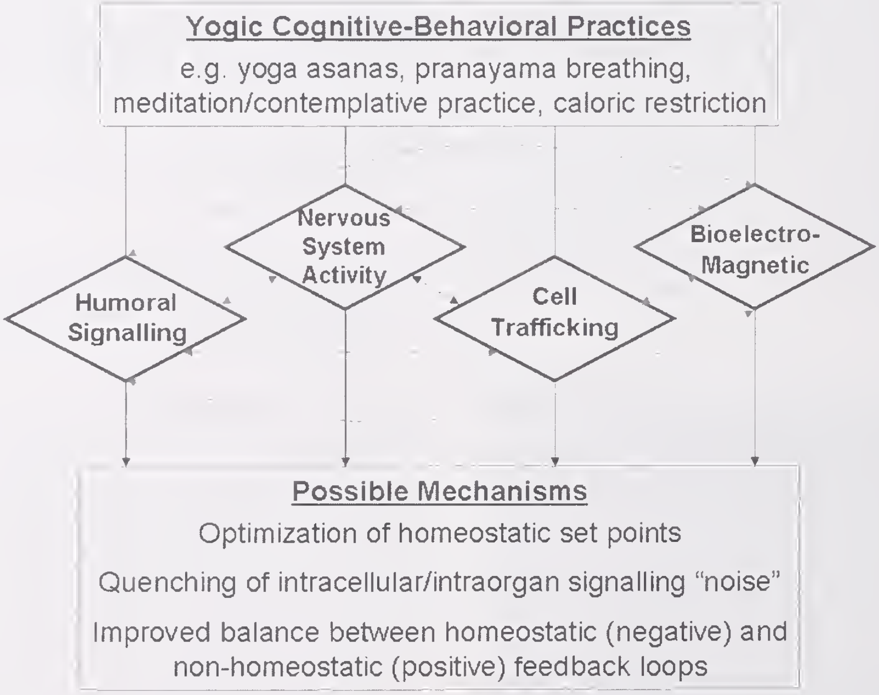

Yogic Cognitive-Behavioral Practices
e.g. yoga asanas, pranayama breathing, meditation/contemplative practice, caloric restriction
Nervous System Activity
Bioelectro-Magnetic
Humoral Signalling
Cell Trafficking
Possible Mechanisms
Optimization of homeostatic set points
Quenching of intracellular/intraorgan signalling "noise"
Improved balance between homeostatic (negative) and non-homeostatic (positive) feedback loops

Fig 1. Yogic cognitive-behavioral practices.

delayed aging, and amelioration of chronic illness and of stress from disability.

Humoral Signaling

Circulating signaling molecules, including hormones, neurotransmitters, growth factors, and the network of cytokines, chemokines, and adipokines (CCA network) link all of the cells and tissues of the body, in health or disease. Although many of these have organ-specific names based on how they were first discovered (eg, hepatocyte growth factor, erythropoietin), they are now recognized as being systemic signaling pathways shared by most, if not all, organs. Thus, if any CBP intervention can modulate these humoral networks on physiological levels, it will have systemwide effects. The largest body of investigation documenting physiologic effects of yogic CBPs, perhaps in part because of the ease of serum sampling and testing, involves such circulating molecules.

A prime example is melatonin. Liou *et al.* have shown, by functional magnetic resonance imaging, that the Chinese meditative practice of "quiet sitting" (similar to mindfulness and Zen "just sitting" meditations) leads to pineal gland activation.[2] It has also been shown that meditation acutely increases circulating melatonin levels in the hours after a period of meditation.[3] In fact, in older adults serum melatonin is significantly elevated toward youthful levels by yoga asana practice, different forms of meditation, and pranayama breathing, both in response to individual sessions of practice (~1 hour) and also as a sustained baseline elevation in long-term practitioners (this is an uncommon study inclusive of long-term practitioners).[4]

Melatonin is most commonly associated with its role in maintenance of circadian rhythms (to which we will return), but a large body of evidence shows that the molecule has a far wider variety of important physiologic effects. A detailed review of these studies is beyond the scope of this article, and has been extensively covered elsewhere.[1,5,6] However, it is worth noting that, being small and amphophilic, the molecule distributes throughout all tissue compartments and fluids and can be found in greatest subcellular concentrations in mitochondria.[7] Its potential impact on perhaps every tissue is therefore profound, leading some to consider it a "master regulatory molecule" not only for the circadian clock, but for other physiologic processes as well.[8]

Melatonin promotes survival and/or regeneration in the liver, skin, hair, bone, hematopoiesis, brain, eyes, heart, gastrointestinal (GI) tract, endothelium, neuroendocrine-reproductive axis, and muscle.[1,5,6] It has been shown to protect healthy tissues from a range of offending agents (toxins, radiation, caloric and mechanical insult) through antioxidant, antiapoptotic, and other mechanisms to stimulate regeneration in healthy tissues.[9,10] It selectively and differentially induces apoptosis in malignant cells (and tissues with other pathologies), including in the breast, lungs, stomach, endometrium, ovaries, cervix, prostate, GI tract, and bone marrow.[9–14]

Pierpoli and colleagues found that both chronic melatonin administration and cross-transplantation of pineal glands from young to old mice lead to a broad range of apparent aging-delaying and/or aging-reversing effects in behavior and physical appearance.[15] Of course, therefore, if yogic CBPs increase pineal activation and circulating melatonin levels, the same beneficial effects found by Pierpoli are likely to result without the need for ingestion of exogenous compound.

Melatonin is not the only humoral signaling molecule that is so affected by yogic CBPs. Dehydroepiandrosterone (DHEA), increasingly known as a cytoprotective, pleiotropic, "ubiquitous" antiaging steroid, has also recently been found to stimulate stem/progenitor cell–based regenerative effects in the brain.[16] The salubrious effects of caloric restriction are attributed, in part, to the maintenance of more youthful levels of DHEA and melatonin, whereas meditation has been found in some studies to increase circulating DHEA to levels more typical of individuals 5–10 years younger on average.[16,17]

Some forms of meditation have also been found to enhance activity of the growth hormone/insulin-like growth factor-I (GH/IGF-1) axis, which is claimed to account for >80% of postnatal growth, and recently enhancement of activity in this axis has been found to promote stem and progenitor cell activation, proliferation, and mobilization.[18] Arginine vasopressin (AVP), circulating levels of which are increased during meditation by approximately 5–7 orders of magnitude,[19] has been found to significantly increase myogenesis and may also play a trophic role in both the peripheral nervous system and the central nervous system (CNS), where it is known to enhance learning and memory through up-regulation of other growth factors.[20–22] Constitutive nitric oxide may also crucially contribute to beneficial therapeutic, regenerative effects in a diverse range of pathologies and exert a "global healing effect"[23] and has been shown to be modulated through Zen meditation practices.[24]

Another very important player in humoral signaling is the CCA network, which is involved in immune reactions of different kinds and magnitude. These reactions are not instantaneous, but take time to occur and propagate through the network and likewise take time to resolve. For example, it takes 1.5–2 hours for the monocytes in a human body to respond to lipopolysaccharide injection.[25] Thus, the resources of this network mobilize in well-defined units over time: increasing concentrations of proinflammatory molecules are followed by increased synthesis of anti-inflammatory molecules to balance the response and eventually resolve the reactive change once the challenge has passed.

However, if not resolved, increased concentrations of proinflammatory and anti-inflammatory molecules are sustained in nonhomeostatic states of imbalance. Evidence suggests that yogic CBPs may inhibit production of such states of imbalance or facilitate resolution back to homeostasis. Indeed, several randomized clinical trials have shown that such CBPs can decrease inflammation (and, not incidentally, thereby improve quality of life and other important, vital parameters in different diseases).

Randomized clinical trials in which New York Heart Association Class I-III heart failure patients were randomized to yoga treatment or standard medical therapy revealed that an 8-week regimen of yoga in addition to standard medical therapy improved exercise tolerance and positively affected levels of inflammatory markers in patients with chronic heart failure. There were statistically significant reductions in serum levels of interleukin-6 (IL-6) and C-reactive protein (CRP) in the yoga-treatment group, and there was also a trend toward improvements in quality of life.[26] Likewise, application of cognitive-behavioral therapy and mindfulness meditation for pain in patients with rheumatoid arthritis promotes actual decrease in mitogen-stimulated levels of IL-6, along with improvement in self-reported pain control and in coping efficacy.[27]

In another example we find evidence that emphasizes not only the utility of a yogic intervention for optimizing health, but confirms that the most benefit is through long-term practice rather than the more commonly studied acute training. As such, it may be seen as a personalized, preventive therapy and of particular import for our contemporary society. Namely, the study at the Institute for Behavioral Medicine Research at The Ohio State University was aimed to investigate the mechanisms underlying hatha yoga's potential stress-reduction benefits. It included 50 healthy women (mean age 41.32 years; range, 30–65 years), 25 novices, and 25 experts (at least 2 years of practice). The results were remarkable: although novices and experts did not differ on key dimensions, including age, abdominal adiposity, and cardiorespiratory fitness, novices' serum IL-6 levels were 41% higher than those of experts across sessions, and the odds of a novice having detectable CRP were 4.75× those of an expert. Thus, in this setting, it is long-term practice, not an acute intervention, that results in beneficial physiologic changes.[28]

Although data about possible connections between CBPs and the CCA network are limited in both numbers of actual randomized trials and in the sets of molecules those trials investigate, it is important to recognize the complexity and diversity of CCA network functions for explanations of possible mechanisms through which CBPs can impact on site-specific molecular and cellular events in the body as well as systemic improvements in aging and/or illness.

Thus, cytokines, chemokines, and adipokines function not only as proinflammatory and anti-inflammatory agents, but also as neurotransmitters, immunosensors and signaling agents. Advances in modern molecular biology, genetics, immunology, and biomedical engineering provide abundant data demonstrating that the immune system is a sophisticated sensory organ capable of detecting and sorting global/systemic information and translating this information into modulations of tissue with site-specific, local effects the way other sensory organs use their own biological mechanisms for detection and sorting of input signals from the environment and then translate these into perception.

In addition to positive selection driving the evolution of cell-to-cell communicators in the mammalian immune system (like gamma chain receptor interleukins IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21),[29] regulation of CCA network takes place on multiple levels of the organism: intracellular and intercellular interactions within specific tissues, between organ systems, and in maintenance of the biological clock. This immune-modulatory network tightly connects all aspects of the body, providing a special biochemical language for bi-directional communications in different situations. For example, members of the tumor necrosis factor and tumor necrosis factor receptor–associated factors family serve as signaling molecules for numbers of physiological processes including adaptive immunity, innate immunity, bone metabolism, and the development of several tissues such as lymph nodes, mammary glands, skin, and the CNS.[30,31]

Moreover, the CCA network is involved in sleep regulation, which brings us back around, of course, to melatonin physiology.[32,33] The CCA network and circadian rhythms modulated, in part, by melatonin are intimately linked.[34] So, not only the median levels of members of the CCA network are important, but circadian variations of their concentrations probably play essential roles in maintenance of normal homeostasis and immune response.[35–39] This reciprocal relationship is demonstrated by the strong association between circadian melatonin production, activation/expression of clock genes (*CRY1* and *PER1*),[40] and adjustments in the immune system parameters.[41] This multilevel physiological rhythmicity depends on, but is not limited by, internal biological clocks and external factors like day/night and seasonal periodicity and is frequently altered during senescence[42] and in different disorders such as cerebrovascular stroke,[43] acute myocardial infarction,[44] and major depressive disorder.[45]

This brief and limited overview might suggest a working hypothesis that yogic CBPs activate the pineal gland, in turn leading to increased levels of circulating melatonin, in turn leading to changes in CCA network signaling. However, such a unidirectional, hierarchical cascade is unlikely. Feedback loops help all members of these interlocked systems to communicate and influence each other. Thus, we have a glimpse of how application of a systems-biologic approach to their understanding will provide for more nuanced understandings. Collaborators from Rutgers University and the University of Medicine and Dentistry, New Jersey-Robert Wood Johnson Medical School, responding to that broadened perspective, have thus developed a mathematical model of the interplay between inflammation and circadian rhythms, based on the hypothesis that these diurnal rhythms are entrained by the cyclic production of the hormones cortisol and melatonin, as stimulated by the central clock in the suprachiasmatic nucleus.[46]

Nervous System Activity

A second transduction pathway involves both CNS activity and signaling by peripheral nerves. Impact of yogic CBPs on CNS activity is, again, beyond the scope of this article (and the expertise of its authors), but has been well summarized elsewhere, in particular for impacts made by meditative practice.[47,48] Here we will for the most part focus on exploring how yogic CBPs can influence human physiology through peripheral nervous system activity.

For example, vagal tone, most often assessed through measurements of heart rate variability (HRV), can be modulated by acupuncture[49,50]; transcendental, Zen and other meditation practices[24,51–54]; and yogic breathing.[24,49,53–57] Intriguingly, Peressutti *et al.* found that not only could Zen meditation alter HRV, but it did so in proportion to the level of experience of the meditator.[58] Influences of such CBP on well-being may be more profound and far-reaching than changes in HRV. Tracey and colleagues have demonstrated that the efferent aspects of the vagal nerve complex play a crucial role in modulating the body's inflammatory response to infectious disease, other forms of physical insult, and psychosocial stress by modulating nuclear factor-kappa B (NF-kappaB)–a key regulator of, among many other pathways, the CCA network–though acetylcholine binding to the alpha-7 nicotinic receptors on activated macrophages and other cytokine-producing cells.[59–61]

In another example, Brown and Gerbarg describe yogic breathing practices, variously including aspects of slow breathing, laryngeal contracture, inspiration against airway resistance, prolonged expiration against airway resistance, and breath-holding, and how these affect physiologic responses to stress as well as maintenance of stabilization of baseline functioning. They do so through mechanisms involving both afferent vagal activity (altering CNS behaviors) and efferent vagal activity, as well as through decreased chemoreflex sensitivity and improved baroreflex sensitivity.[49,50,56] Again, as with topics described above, a detailed review of these findings is beyond the scope of this article, but one may summarize a large body of experimental literature by indicating that vagal nerve afferents synapse in the medulla on the nucleus tractus solitarius, which in turn communicates via the parabrachial nucleus to the thalamic nuclei (leading to modulatory effects on frontal and parietal occipital cortices) and the limbic system. Through the limbic system, directly or indirectly through other centers, neuroendocrine outputs are then modulated, including reduction of cortisol and increases in AVP, oxytocin, and prolactin.

These last links in the chain of Brown and Gerbarg's neurophysiologic model of pranayama are yet another indication of the interconnectedness of these transduction pathways for the impact of yogic CBPs on physiologic functions. Along these lines, we note that the neuroendocrine and immune systems produce a common set of peptide and nonpeptide neurotransmitters and cytokines that act on a common repertoire of receptors in the 2 systems. This complete biochemical information circuit between neurons and immune cells allows the immune system to function as a sensory organ,[62] providing an evidence of interconnection of 2 transduction pathways: humoral factors and nerve signaling. For example, it has been recently reported that oxytocin, a primitive neurohypophyseal hormone considered a modulator of lactation and social bonding, is a direct regulator of bone mass.[63] Additionally, depending on diet composition, certain cytokines are expressed in the hypothalamus, contributing to the activation of intracellular inflammatory signal transduction,[64] and calorie restriction may provide neuroprotection to the aging brain by preserving DNA-repair enzymes in their intact form and/or up-regulating specific antiapoptotic proteins involved in neuronal cell death.[65]

Cell Trafficking

We have already noted above that Tracey's work indicates the possibilities for vagal innervation of circulating immunocytes, such as macrophages, through stimulation of nicotinic receptors on their surface. This brings us to the third transduction pathway of cell trafficking. The field of psychoneuroimmunology, for which Tracey's efforts represent an important component, details how circulating cells such as macrophages and lymphocytes are not independent of external or internal environmental influences and their activities are mediated through CNS and peripheral nervous system actions. As yogic CBPs can effect CNS and at least vagal functioning, so may they in turn effect mobilization, activation states, and cell production.

There is comparatively little data demonstrating effects of yogic CBPs on immunocyte trafficking, however. Two recent examples are fairly typical and straightforward. Kochupillai *et al.* show that cancer patients taught yogic breathing practices increased natural killer cells significantly ($P < 0.001$) at 12 and 24 weeks of the practice compared with baseline, and the increase in natural killer (NK) cells at 24 weeks was significant ($P < 0.05$) compared with controls.[51] Witek-Janusek and colleagues studied women recently diagnosed with breast cancer who were enrolled in a mindfulness-based stress reduction (MBSR) program. These patients, over time, re-established their homeostatic NK cell activity and cytokine levels. In contrast, breast cancer patients in the non-MBSR control group exhibited continued reductions in NK cell activity and interferon-gamma production, with increased production of cytokines IL-4, IL-6 (proinflammatory), and IL-10 (anti-inflammatory).[66]

That immunocytes traffic through the body as part of immune system defense is obvious, though the diversity of effects ranges far beyond the most traditionally studied antimicrobial or antitumor defenses. Immunocytes are increasingly recognized for their importance in promoting cell repopulation, revascularization, and modulation of scar formation. In our own work and that of other colleagues with expertise in a variety of organs, the role of circulating stem/progenitor cells has been a central focus, with marrow-derived, hematopoietic, and mesenchymal cells traveling throughout the body to contribute to repair at sites of injury.[67] One possible route for such an effect (and yet another indication of how intimately interconnected are all these "separate" transduction pathways) is that the CNS can directly modulate hematopoietic stem cell activity in the marrow.[68,69] We speculate that these signals may be influenced by yogic practices as well.

Indeed, sites of injury themselves produce humoral signals that further influence cell trafficking.[70] Intriguingly, the work of Ellen Heber-Katz of the Wistar Institute,[71,72] showing scarless healing in mixed lymphocyte reaction mice mediated by circulating monocytes homing to sites of wounding, raises questions about well-documented bloodless and scarless healing in response to acute piercing and cutting in rituals performed during profound meditative and trance-induced mental states.[72,73]

Bioelectromagnetism

This area is the most difficult to conceptualize, in that the scientific as well as daily languages available for its description are profoundly limited.[74] However, for the purposes of these efforts, we will focus on electrical conductivity reflective of ionic or electron flux. Obviously, some aspects of electrical flux in the body are also mediated by nerves, but there is electrical conduction through all tissues and it plays significant roles in diverse cellular-level and tissue-level processes.[75]

An interesting example is that of skin, in which electrical conduction plays a critical role in tissue repair and regeneration, with specific gene expression playing a (necessary) role in these basic wound-healing mechanisms.[76] There is a normal, actively maintained voltage drop from the epidermal basement membrane to the skin surface. Disruption of epidermis by wounding shifts the gradient toward the wound. Ex vivo studies of cultured squamous cells showed that the direction of this gradient guides cell migration into the wound, and ex vivo and in vivo studies demonstrated that increasing the strength of local electrical fields in the range of the endogenous wound electrical field enhanced healing. Finally, these investigators showed that such enhancement was abolished by disruption of genetic pathways encoding for phosphatidlyinositol-3-OH kinase-gamma.

When combined with research that shows acupuncture or application of "energy" by Chi masters can increase skin conductivity or that maintenance of different meditative absorptive states can do likewise,[77–79] the possibility that yogic practices might directly impact on rate of at least cutaneous wound healing may be significant.[74] However, as stated, this represents a considerably understudied area.

TOWARD A CHANGING PARADIGM FOR INVESTIGATION

Although these examples demonstrate impacts on various signaling pathways, sometimes in great detail and sometimes with obvious potential to maximize repair after injury, improve restoration of normal baselines following physiologic stress, and even promote rejuvenation in the face of aging, they do not actually answer the question of how alteration of pathways leads to systemic improvements. Indeed, most studies to date remain locked in a reductionist paradigm of investigation, focused on very small numbers of possible effects with very little to suggest a holistic consideration of the body as an extraordinarily complex system of interacting parts, a whole that is far beyond explanations supplied by assembled pieces. Thus, as we suggested in discussing of melatonin and its role in transducing effects of yogic CBPs, although the data support the notion of a hierarchical, top-down flow of effects, in fact this would be an oversimplification that misses the greater, complex interweaving of possibilities. In the discussion of the impact on CNS behavior and subsequent modulation of cortisol, oxytocin, prolactin, and AVP, might it not also impact on the pineal gland and on melatonin production? Thus, with consideration of that possibility alone, the hope for a unidirectional, hierarchical mode explanation of physiologic effects of yogic CBPs crumbles.

We suggest, as have others,[51] that a systems approach would be not only beneficial, but necessary for a full understanding of these issues. The work of Scheff *et al.*, noted earlier, producing mathematical models for the relationships of circadian rhythms and CCA network signaling, is an example of this methodology for understanding complex biological systems and the impacts of interventions on their behaviors. The data necessary for such model-building would ideally be derived, of course, from high-throughput, real-time analyses of such domains as functional genomics, proteomics, and metabolomics. It is in the changing patterns of the whole system that the profoundest understanding of yogic mechanisms will be reached. It is also with such an approach that we are more likely to understand and appreciate the benefits of their consistent, long-term practice rather than their acutely applied, therapeutic use. After all, it is unlikely that long-term use is simply going to achieve the summation of cumulative, single "doses." On that basis, we propose 3 possible conceptual ways of considering this larger question of mechanism that implies the need for different kinds of investigations than are currently typical, investigations that pursue an understanding of how the biological system of the body, as a whole, is modulated by yogic CBPs (Figure 1).

The first relates to the fact that in response to stress, resulting in or arising from aging, disease, injury, or disability, there are changes to physiologic setpoints. Biological systems are rarely if ever truly stable. Rather, they self-regulate within a homeostatic range. Disturbances of the normal, baseline state of the organism may result in alterations of the setpoints that define and bound such homeostatic control. Normal body temperature, for example, follows a circadian rhythm. In response to viral infection, fever may be an appropriate response in which the setpoints for homeostatic control are shifted upward to facilitate antiviral responses; however, circadian rhythms are maintained, only the setpoints for temperature control for the organism have changed. Thus, we propose that yogic practices may help to restore physiologic setpoints to normal.

We propose that yogic practices may help to restore physiologic setpoints to normal.

Additional hypotheses follow upon consideration that bodies represent examples of complex adaptive systems; ie, they are emergent, self-organized structures arising from the interactions of their component cellular and molecular parts. As we have discussed in greater detail elsewhere, the cells and molecules of the body give rise to the development of the body over time (from embryo to fetus to postnatal life) and also to its nature at any given moment within its lifespan.[52,80–82] They do so because they fulfill certain criteria: they are sufficiently numerous; all elements within the system are responding only to local signals, without global sensing; there is an overall balance of negative over positive feedback loops; and there is quenched disorder in the system. These last 2 criteria suggest additional hypotheses.

The need for an overall balance of negative, or homeostatic, feedback loops for adaptive stability of a living system is fairly self-evident. Complexity theory describes that self-organization can still occur with a balance of positive feedback loops, but rather than being adaptive, these forms of self-organization are energy expending and self-limiting. Tornadoes and hurricanes are examples; cancer would be another, more germane to this current context. The emergence of cancer may be thought of as the gradual dominance of positive feedback signaling between cancer cells and normal cells/matrix that leads to the expanding, energy-expending, ultimately self-limiting growth of malignancy. Thus, we hypothesize that yogic CBPs may inhibit emergence of positive feedback loops or enhance homeostatic negative feedback.

Yogic CBPs may inhibit emergence of positive feedback loops or enhance homeostatic negative feedback.

The oncostatic properties of melatonin or the inhibition of NF-kappaB in macrophages by vagal stimulation are examples of suppressive control of positive feedbacks. Likewise, pranayama enhancements of baroreflexes and increasing numbers of circulating natural killer cells in breast cancer patients after mindfulness meditation may be demonstrations of enhanced negative, homeostatic feedback loops.

The presence of quenched disorder in complex systems is important, because without it there would be no ability for the system to explore adaptive responses to a changing environment.[81,82] Too little disorder in the system and there would be no opportunity for the system to change, leading to eventual collapse. On the other hand, too much disorder and the system will not be able to respond efficiently to the challenge of a changing environment, also leading to collapse. Sufficiently large increases in disorder cannot be overcome by yogic CBPs (or any other intervention, for that matter). However, subtly increased "noise" in the system that is not immediately fatal may be a way of conceiving some aspects of aging, chronic disease, injury, and disability.

For example, a majority of molecules in the CCA network are present in very small serum/plasma concentrations, only several picograms per milliliter; thus, seemingly negligible variations might have dramatic impact on function. For example, whereas normal serum albumin level is 30–50 g/L, serum levels of one of the most studied proinflammatory cytokines, IL-6, are approximately 1 pg/mL, or 0.000000000001 g/L (reported values of IL-6 may vary depending on measurement methods). The importance of such pg/mL values can be illustrated, for example, by the results of the study of a possible connection of serum IL-6 level and the development of disability in persons aged 71 years or older, where even after adjustments for multiple confounders there was strong nonlinear association between the increased risk of mobility/disability and IL-6 concentration, with the risk rising rapidly beyond plasma levels of only 2.5 pg/mL.[83] Additionally, every particular change should be relatively big in order to be noticeable and, therefore, the system as a whole should be relatively free of "noise" to notice subtle but important changes. There should also be sufficient numbers of active receptors to identify such subtle changes and initiate proper response.

Thus, low-level, disordered signaling, promoted through aging, disability, or disease, may swamp the delicate signaling required for maintenance of health and optimization of the organism. Chronic infections, dysregulated gene activation in aged cells, and exposures to intermittent or persistent environmental stresses are examples of challenges to the system that might not decisively interfere with functioning, but will lead to sufficient, random fluctuations or "noise," such that normative, homeostatic signaling is impaired. Therefore, a third proposal is that "noise" in the signaling networks may interfere with physiologic homeostasis and that yogic CPBs may assist in quenching such disorder.

"Noise" in the signaling networks may interfere with physiologic homeostasis and yogic (CBPs) may assist in quenching such disorder.

CONCLUSION

It would be disingenuous at this point in time for a critic or skeptic to declare that there is insufficient evidence with regard to the possible beneficial effects of various yogic practices for optimization of health, delay of senescence, and promotion of recovery from disease or injury. As we found it necessary to repeat in several sections of this article, the truly extensive nature of the data in support of these possibilities makes it impossible to contain within a single, short review such as this. Indeed, consideration of a comprehensive textbook may be the next step in moving this field of investigation forward.

However, review of the abundant data available allows for certain principles to emerge with some degree of clarity. First, the physiologic effects of yogic

CBPs can be described as modulating physiologic systems through 4 transduction pathways: humoral signaling, nervous system activity, cell trafficking, and bioelectromagnetism. Some of these (humoral signaling) have been extensively studied, others (cell trafficking and bioelectromagnetism) have not. Clearly, there is much work to be done in all these areas.

At the same time, there is enough data to know that studies of a traditional nature will only yield very limited windows on the mechanisms whereby yogic CBPs can benefit practitioners. After all, it becomes very quickly clear that behaviors within one transduction pathway will rapidly propagate through and effect others, the web of interactions becoming rapidly too complex to contain within a simple framework. Moreover, the overwhelming numbers of studies are of short-term duration in subjects who have only recently been taught the intervention.

Furthermore, the models of such mechanisms derived from such narrowly focused studies have relatively little explanatory power. The example we gave of melatonin physiology is emblematic: yogic practices can activate the pineal gland, presumably leading to the confirmed increases in circulating melatonin, which in turn leads to tissue-protective effects through its antioxidant and oncostatic actions. However, could one not simply take a little bit of melatonin at bedtime and achieve these same effects? But the actions of yogic practices are more subtle and are not strictly dose dependent; practitioners readily claim that long-term practice promotes greater benefits than could be explained by single, repeated "dosing" of an intervention. Where is there room in the reductionist approach to begin to assess the effects of long-term practice?

To address these possibilities, a systems-biology approach must be applied to the study of the physiologic effects of yogic CBPs. In doing so, we will develop a richer set of concepts to understand these practices, particularly as they deepen in long-term practice; a new set of hypotheses to guide more subtle (and more ambitious) study designs; and the possibility of devising ways to directly compare one practice to another.

We suggest 3 such hypotheses, in part based on our understanding of the body as a complex adaptive system, namely that yogic practices may: (1) promote restoration of physiologic setpoints to normal after derangements secondary to disease or injury, (2) promote homeostatic negative feedback loops over nonhomeostatic positive feedback loops in molecular and cellular interactions, and (3) quench abnormal "noise" in cellular and molecular signaling networks arising from environmental or internal stresses. These hypotheses are not in opposition and may all function in parallel or in series in different circumstances and at different times. They are also merely the first 3 that we have imagined—there may indeed be many more. We offer these as places to start.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Bushell WC, Spector NH, Thiese ND. From the global to the local: possible pathways for the transduction of Indo-Sino-Tibetan cognitive-behavioral practices into site-specific, tissue-regenerative effects. *Ann N Y Acad Sci* 2009; 1172: 74–87.
2. Liou CH, Hsieh CW, Hsieh CH, et al. Studies of Chinese original quiet sitting by using functional magnetic resonance imaging. *Conf Proc IEEE Eng Med Biol Soc* 2005; 5: 5317–5319.
3. Tooley GA, Armstrong SM, Norman TR, Sali A. Acute increases in night-time plasma melatonin levels following a period of meditation. *Biol Psychol* 2000; 53: 69–78.
4. Martarelli D, Cocchioni M, Scuri S, Pompei P. Diaphragmatic breathing reduces exercise-induced oxidative stress. *Evid Based Complement Alternat Med* 2009; doi:10.1093/ecam/nep169.
5. Bonnefont-Rousselot D, Collin F. Melatonin: action as antioxidant and potential applications in human disease and aging. *Toxicology* 2010; doi:10.1016/j.tox.2010.04.008.
6. Bushell WC. From molecular biology to anti-aging cognitive-behavioral practices: the pioneering research of Walter Pierpaoli on the pineal and bone marrow foreshadows the contemporary revolution in stem cell and regenerative biology. *Ann N Y Acad Sci* 2005; 1057: 28–49.
7. Paradies G, Petrosillo G, Paradies V, et al. Melatonin, cardiolipin and mitochondrial bioenergetics in health and disease. *J Pineal Res* 2010; 48: 297–310.
8. Sánchez-Barceló EJ, Mediavilla MD, Tan DX, Reiter RJ. Clinical uses of melatonin: evaluation of human trials. *Curr Med Chem* 2010; 17: 2070–2095.
9. Reiter RJ, Tan DX, Fuentes-Broto L. Melatonin: a multitasking molecule. *Prog Brain Res* 2010; 181: 127–151.
10. Pandi-Perumal SR, Srinivasan V, Maestroni GJ, et al. Melatonin: Nature's most versatile biological signal? *FEBS J* 2006; 273: 2813–2838.
11. Schernhammer ES, Schulmeister K. Melatonin and cancer risk: does light at night compromise physiologic cancer protection by lowering serum melatonin levels? *Br J Cancer* 2004; 90: 941–943.
12. Sainz RM, Mayo JC, Rodriguez C, et al. Melatonin and cell death: differential actions on apoptosis in normal and cancer cells. *Cell Mol Life Sci* 2003; 60: 1407–1426.
13. Wölfler A, Caluba HC, Abuja PM, et al. Prooxidant activity of melatonin promotes fas-induced cell death in human leukemic Jurkat cells. *FEBS Lett* 2001; 502: 127–131.

14. Karasek M. Melatonin, human aging, and age-related diseases. *Exp Gerontol* 2004; 39: 1723–1729.
15. Pierpaoli W, Bulian D, Dall' Ara A, et al. Circadian melatonin and young-to-old pineal grafting postpone aging and maintain juvenile conditions of reproductive functions in mice and rats. *Exp Gerontol* 1997; 32: 587–602.
16. Carlson LE, Speca M, Patel KD, Goodey E. Mindfulness-based stress reduction in relation to quality of life, mood, symptoms of stress and levels of cortisol, dehydroepiandrosterone sulfate (DHEAS) and melatonin in breast and prostate cancer outpatients. *Psychoneuroendocrinology* 2004; 29: 448–474.
17. Vera FM, Manzaneque JM, Maldonado EF, et al. Subjective sleep quality and hormonal modulation in long-term yoga practitioners. *Biol Psychol* 2009; 81: 164–168.
18. Wahbeh H, Calabrese C, Zwickey H. Binaural beat technology in humans: a pilot study to assess psychologic and physiologic effects. *J Altern Complement Med* 2007; 13: 25–32.
19. O' Halloran JP, Jevning R, Wilson AF, et al. Hormonal control in a state of decreased activation: potentiation of arginine vasopressin secretion. *Physiol Behav* 1985; 35: 591–595.
20. Holsboer F, Ising M. Stress hormone regulation: biological role and translation into therapy. *Annu Rev Psychol* 2010; 61: 81–109, C1–11.
21. Reghunandanan V, Reghunandanan R, Mahajan KK. Arginine vasopressin as a neurotransmitter in brain. *Indian J Exp Biol* 1998; 36: 635–643.
22. Alonso G. Vasopressin and angiogenesis [in French]. *J Soc Biol* 2009; 203: 39–47.
23. Stefano GB, Esch T. Integrative medical therapy: examination of meditation's therapeutic and global medicinal outcomes via nitric oxide (review). *Int J Mol Med* 2005; 16: 621–630.
24. Kim DH, Moon YS, Kim HS, et al. Effect of Zen Meditation on serum nitric oxide activity and lipid peroxidation. *Prog Neuropsychopharmacol Biol Psychiatry* 2005; 29: 327–331.
25. Haimovich B, Calvano J, Haimovich AD, et al. In vivo endotoxin synchronizes and suppresses clock gene expression in human peripheral blood leukocytes. *Crit Care Med* 2010; 38: 751–758.
26. Pullen PR, Nagamia SH, Mehta PK, et al. Effects of yoga on inflammation and exercise capacity in patients with chronic heart failure. *J Card Fail* 2008; 14: 407–413.
27. Zautra AJ, Davis MC, Reich JW, et al. Comparison of cognitive behavioral and mindfulness meditation interventions on adaptation to rheumatoid arthritis for patients with and without history of recurrent depression. *J Consult Clin Psychol* 2008; 76: 408–421.
28. Kiecolt-Glaser JK, Christian L, Preston H, et al. Stress, inflammation, and yoga practice. *Psychosom Med* 2010; 72: 113–121.
29. O'Connell MJ, McInerney JO. Gamma chain receptor interleukins: evidence for positive selection driving the evolution of cell-to-cell communicators in the mammalian immune system. *J Mol Evol* 2005; 61: 608–619.
30. Wu H, Arron JR. TRAF6, a molecular bridge spanning adaptive immunity, innate immunity and osteoimmunology. *Bioessays* 2003; 25: 1096–1105.
31. Nakamura I, Kadono Y, Takayanagi H, et al. IL-1 regulates cytoskeletal organization in osteoclasts via TNF receptor-associated factor 6/c-Src complex. *J Immunol* 2002; 168: 5103–5109.
32. Kubota T, Brown RA, Fang J, Krueger JM. Interleukin-15 and interleukin-2 enhance non-REM sleep in rabbits. *Am J Physiol Regul Integr Comp Physiol* 2001; 281: R1004–R1012.
33. Krueger JM, Obál FJ, Fang J, et al. The role of cytokines in physiological sleep regulation. *Ann N Y Acad Sci* 2001; 933: 211–221.
34. Lissoni P. The pineal gland as a central regulator of cytokine network. *Neuro Endocrinol Lett* 1999; 20: 343–349.
35. Vgontzas AN, Bixler EO, Lin HM, et al. IL-6 and its circadian secretion in humans. *Neuroimmunomodulation* 2005; 12: 131–140.
36. Vgontzas AN, Papanicolaou DA, Bixler EO, et al. Circadian interleukin-6 secretion and quantity and depth of sleep. *J Clin Endocrinol Metab* 1999; 84: 2603–2607.
37. Redwine L, Hauger RL, Gillin JC, Irwin M. Effects of sleep and sleep deprivation on interleukin-6, growth hormone, cortisol, and melatonin levels in humans. *J Clin Endocrinol Metab* 2000; 85: 3597–3603.
38. Vgontzas AN, Zoumakis E, Bixler EO, et al. Adverse effects of modest sleep restriction on sleepiness, performance, and inflammatory cytokines. *J Clin Endocrinol Metab* 2004; 89: 2119–2126.
39. Vgontzas AN, Zoumakis M, Bixler EO, et al. Impaired nighttime sleep in healthy old versus young adults is associated with elevated plasma interleukin-6 and cortisol levels: physiologic and therapeutic implications. *J Clin Endocrinol Metab* 2003; 88: 2087–2095.
40. Lincoln GA. Decoding the nightly melatonin signal through circadian clockwork. *Mol Cell Endocrinol* 2006; 252: 69–73.
41. Carrillo-Vico A, Guerrero JM, Lardone PJ, Reiter RJ. A review of the multiple actions of melatonin on the immune system. *Endocrine* 2005; 27: 189–200.
42. Sadki A, Bentivoglio M, Kristensson K, Nygard M. Suppressors, receptors and effects of cytokines on the aging mouse biological clock. *Neurobiol Aging* 2007; 28: 296–305.
43. Szczudlik A, Dziedzic T, Bartus S, et al. Serum interleukin-6 predicts cortisol release in acute stroke patients. *J Endocrinol Invest* 2004; 27: 37–41.
44. Dominguez-Rodriguez A, Abreu-Gonzalez P, Garcia M, et al. Light/dark patterns of interleukin-6 in relation to the pineal hormone melatonin in patients with acute myocardial infarction. *Cytokine* 2004; 26: 89–93.
45. Alesci S, Martinez PE, Kelkar S, et al. Major depression is associated with significant diurnal elevations in plasma interleukin-6 levels, a shift of its circadian rhythm, and loss of physiological complexity in its secretion: clinical implications. *J Clin Endocrinol Metab* 2005; 90: 2522–2530.
46. Scheff JD, Calvano SE, Lowry SF, Androulakis IP. Modeling the influence of circadian rhythms on the acute inflammatory response. *J Theor Biol* 2010; 264: 1068–1076.
47. Lutz A, Slagter HA, Rawlings NB, et al. Mental training enhances attentional stability: neural and behavioral evidence. *J Neurosci* 2009; 29: 13418–13427.
48. Brefczynski-Lewis JA, Lutz A, Schaefer HS, et al. Neural correlates of attentional expertise in long-term

meditation practitioners. *Proc Natl Acad Sci U S A* 2007; 104: 11483–11488.

49. Brown RP, Gerbarg PL. Sudarshan Kriya yogic breathing in the treatment of stress, anxiety, and depression: part II–clinical applications and guidelines. *J Altern Complement Med* 2005; 11: 711–717.
50. Brown RP, Gerbarg PL. Sudarshan Kriya yogic breathing in the treatment of stress, anxiety, and depression: part I–neurophysiologic model [published correction appears in *J Altern Complement Med* 2005;11:383–384]. *J Altern Complement Med* 2005; 11: 189–201.
51. Kochupillai V, Kumar P, Singh D, et al. Effect of rhythmic breathing (Sudarshan Kriya and Pranayam) on immune functions and tobacco addiction. *Ann N Y Acad Sci* 2005; 1056: 242–252.
52. Theise ND. Beyond cell doctrine: complexity theory informs alternate models of the body for cross-cultural dialogue. *Ann N Y Acad Sci* 2009; 1172: 263–269.
53. Phongsuphap S, Pongsupap Y, Chandanamattha P, Lursinsap C. Changes in heart rate variability during concentration meditation. *Int J Cardiol* 2008; 130: 481–484.
54. Wu SD, Lo PC. Inward-attention meditation increases parasympathetic activity: a study based on heart rate variability. *Biomed Res* 2008; 29: 245–250.
55. Pramanik T, Sharma HO, Mishra S, et al. Immediate effect of slow pace bhastrika pranayama on blood pressure and heart rate. *J Altern Complement Med* 2009; 15: 293–295.
56. Brown RP, Gerbarg PL. Yoga breathing, meditation, and longevity. *Ann N Y Acad Sci* 2009; 1172: 54–62.
57. Khattab K, Khattab AA, Ortak J, et al. Iyengar yoga increases cardiac parasympathetic nervous modulation among healthy yoga practitioners. *Evid Based Complement Alternat Med* 2007; 4: 511–517.
58. Peressutti C, Martín-González JM, García-Manso JM, et al. Heart rate dynamics in different levels of Zen meditation. *Int J Cardiol* 2009; doi:10.1016/j.ijcard.2009.06.058.
59. Tracey KJ. Physiology and immunology of the cholinergic antiinflammatory pathway. *J Clin Invest* 2007; 117: 289–296.
60. Goldstein RS, Bruchfeld A, Yang L, et al. Cholinergic anti-inflammatory pathway activity and High Mobility Group Box-1 (HMGB1) serum levels in patients with rheumatoid arthritis. *Mol Med* 2007; 13: 210–215.
61. Rosas-Ballina M, Tracey KJ. The neurology of the immune system: neural reflexes regulate immunity. *Neuron* 2009; 64: 28–32.
62. Blalock JE. The immune system as the sixth sense. *J Intern Med* 2005; 257: 126–138.
63. Tamma R, Colaianni G, Zhu LL, et al. Oxytocin is an anabolic bone hormone. *Proc Natl Acad Sci U S A* 2009; 106: 7149–7154.
64. Velloso LA, Araújo EP, de Souza CT. Diet-induced inflammation of the hypothalamus in obesity. *Neuroimmunomodulation* 2008; 15: 189–193.
65. Hiona A, Leeuwenburgh C. Effects of age and caloric restriction on brain neuronal cell death/survival. *Ann N Y Acad Sci* 2004; 1019: 96–105.
66. Witek-Janusek L, Albuquerque K, Chroniak KR, et al. Effect of mindfulness based stress reduction on immune function, quality of life and coping in women newly diagnosed with early stage breast cancer. *Brain Behav Immun* 2008; 22: 969–981.
67. Krause DS, Thiese ND, Collector MI, et al. Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell* 2001; 105: 369–377.
68. Larsson J, Scadden D. Nervous activity in a stem cell niche. *Cell* 2006; 124: 253–255.
69. Bhatt R, Rameshwar P, Goldstein K, Siegel A. Effects of kindled seizures upon hematopoiesis in rats. *Epilepsy Res* 2003; 54: 209–219.
70. Swenson ES, Kuwahara R, Krause DS, Thiese ND. Physiological variations of stem cell factor and stromal-derived factor-1 in murine models of liver injury and regeneration. *Liver Int* 2008; 28: 308–318.
71. Blankenhorn EP, Bryan G, Kossenkov AV, et al. Genetic loci that regulate healing and regeneration in LG/J and SM/J mice. *Mamm Genome* 2009; 20: 720–733.
72. Heber-Katz E, Gourevitch D. The relationship between inflammation and regeneration in the MRL mouse: potential relevance for putative human regenerative (scarless wound healing) capacities? *Ann N Y Acad Sci* 2009; 1172: 110–114.
73. Jackson JE. The cross-cultural evidence on "extreme behaviors" : what can it tell us? *Ann N Y Acad Sci* 2009; 1172: 270–277.
74. Rosch PJ. Bioelectromagnetic and subtle energy medicine: the interface between mind and matter. *Ann N Y Acad Sci* 2009; 1172: 297–311.
75. McCaig CD, Song B, Rajnicek AM. Electrical dimensions in cell science. *J Cell Sci* 2009; 122: 4267–4276.
76. Zhao M, Song B, Pu J, et al. Electrical signals control wound healing through phosphatidylinositol-3-OH kinase-gamma and PTEN. *Nature* 2006; 442: 457–460.
77. Lee MS, Hong SS, Lim HJ, et al. Retrospective survey on therapeutic efficacy of Qigong in Korea. *Am J Chin Med* 2003; 31: 809–815.
78. Leung AY, Park J, Schulteis G, et al. The electrophysiology of de qi sensations. *J Altern Complement Med* 2006; 12: 743–750.
79. Tang YY, Ma Y, Fan Y, et al. Central and autonomic nervous system interaction is altered by short-term meditation. *Proc Natl Acad Sci U S A* 2009; 106: 8865–8870.
80. Theise ND. Now you see it, now you don't. *Nature* 2005; 435: 1165.
81. Theise ND, d'Inverno M. Understanding cell lineages as complex adaptive systems. *Blood Cells Mol Dis* 2004; 32: 17–20.
82. Theise ND. Implications of 'postmodern biology' for pathology: the Cell Doctrine. *Lab Invest* 2006; 86: 335–344.
83. Ferrucci L, Harris TB, Guralnik JM, et al. Serum IL-6 level and the development of disability in older persons. *J Am Geriatr Soc* 1999; 47: 639–646.

HEED

HEALTH ECONOMIC EVALUATIONS DATABASE

Ensures you make informed decisions about medicines and other healthcare interventions

HEED, the **Health Economics Evaluations Database,** equips health professionals with the tools they need to make informed decisions about medicines and other healthcare interventions. It does so by selecting and reviewing published economic studies which provide comparative analysis of costs and consequences.

Key Features of HEED:

- **Comprehensive coverage** – over 42,000 record (c. 21,000 full reviews and 21,000 bibliographic details only) dating from 1958-2010 with over 5,000 journals searched every week via Medline or EMBASE.
- **Up-to-date** – 250 records added each month.
- **Current** – over 300 articles published in 2010 were reviewed on HEED in the first 4 months of 2010.
- **International** – over 1,200 non-English language articles referenced.
- **Expert** – c. 1,400 non-English articles referenced.
- **Structured** – easily searchable format covering all types of economic evaluation, technology assessed, cost data, ICD codes, ATC Classifications, controlled keywords and a summary of key results.
- **Broad-range of reviews** - including cost-effectiveness analyses, cost-utility analyses, cost benefit analyses, methodology papers and many more.
- **Reliable** – the HEED selection process is a rigorous and exacting procedure that produces comprehensive coverage of economic evaluations.

WILEY-BLACKWELL

How to Order

HEED is available online on a subscription basis from Wiley-Blackwell.

For more information or to request a quote contact your Wiley Account Manager or **email heed@wiley.co.uk**.

http://heed.wiley.com

13542

WILEY-
BLACKWELL

Corporate Sales

Dedicated to supporting your communication needs

Advertising | Books | Reprints
Supplements and Sponsorship

- Through our extensive scientific, clinical and professional publications, we are able to help you reach key markets by developing customized communication plans to support your business
- We offer efficient and knowledgeable service, competitive pricing, dedicated project management and the flexibility to meet the most demanding requirements across our range of products and services
- Whether you are looking for global or localized campaigns, for strategic or tactical support, our publishing teams are knowledgeable and easy to reach in your local time zone

Find out more:
W: **www.wiley.com/wiley-blackwell** | E: **corporatesalesusa@wiley.com**

Europe: Chichester • Oxford • Berlin • Weinheim • Copenhagen • Madrid
North America: Hoboken • Malden | **Australia:** Melbourne | **Asia:** Tokyo • New Delhi • Shanghai • Singapore

The Evidence-Based Resource Trusted by Health Professionals

ESSENTIAL
EVIDENCE
PLUS

Essential Evidence Plus (EE+) is the only point-of-care resource that provides you with the Bottom Line to help your patients. This powerful new e-reference tool includes medical evidence ratings so you know you are getting the best EBM answers for teaching and patient care.

EE+ gives you everything you need to make critical patient care decisions, and includes must-have content, calculators, daily POEM updates, guidelines, and tools that clinicians can rely on for the best answer. It has the latest medical findings pre-screened for validity and relevance to clinical practice, and can be used to search all of the evidence, all of the time, in an instant. **EE+** is packed with great features you won't find with any other comparable product and is the most complete evidence-based medicine resource available!

Unique features include:

- **NEW RESOURCE!** *Essential Evidence* featuring 700+ current clinical topics designed to provide evidence-based answers to important questions concerning symptoms, diseases, and treatment that arise in primary care and hospital practice
- 1,900+ diagnostic test calculators
- 1,700+ history and physical test calculators
- 300+ interactive decision support tools and clinical decision rules
- Daily POEM e-mail alerts to help you stay abreast of the latest and most relevant medical literature and 3,500+ Archived POEMs
- 3,300+ Cochrane Abstracts
- EBM Guidelines brought to you by the Finnish Medical Society with 1,000+ practice guidelines, 3,000+ Evidence-Graded Summaries and 950+ high quality photographs
- Derm Expert viewer with 950+ pictures designed to help clinicians diagnose dermatological conditions
- 1,600+ Guideline Summaries
- 1,500+ Commonly used ICD-9 Codes
- Weekly Podcasts
- CME credit availability for qualified searches

Don't take our word for it. Check out what others are saying about Essential Evidence Plus!

> "**Essential Evidence Plus** is an important tool that will provide primary care clinicians with access to the most extensive source of evidence-based medical information currently available for use at the point of care."
>
> — Michael Wilkes, MD,
> Professor of Medicine and
> Vice Dean of Education,
> University of California, Davis

> "This database will revolutionize the way librarians work with clinicians. [It] has set the standard for all clinical database designers to surpass if they want to succeed."
>
> — Rick Wallace, Outreach Librarian,
> Eastern Tennessee State University

> "...places highly valid, highly relevant information 'at the fingertips' of clinicians while they practice."
>
> — Journal of the Medical
> Library Association

> "This is the most robust, able medical resource available."
>
> — David Stockwell, MD,
> Children's National Medical Center

Sign up for a free trial or subscribe today!
Group price quotations are also available upon request.

www.essentialevidenceplus.com

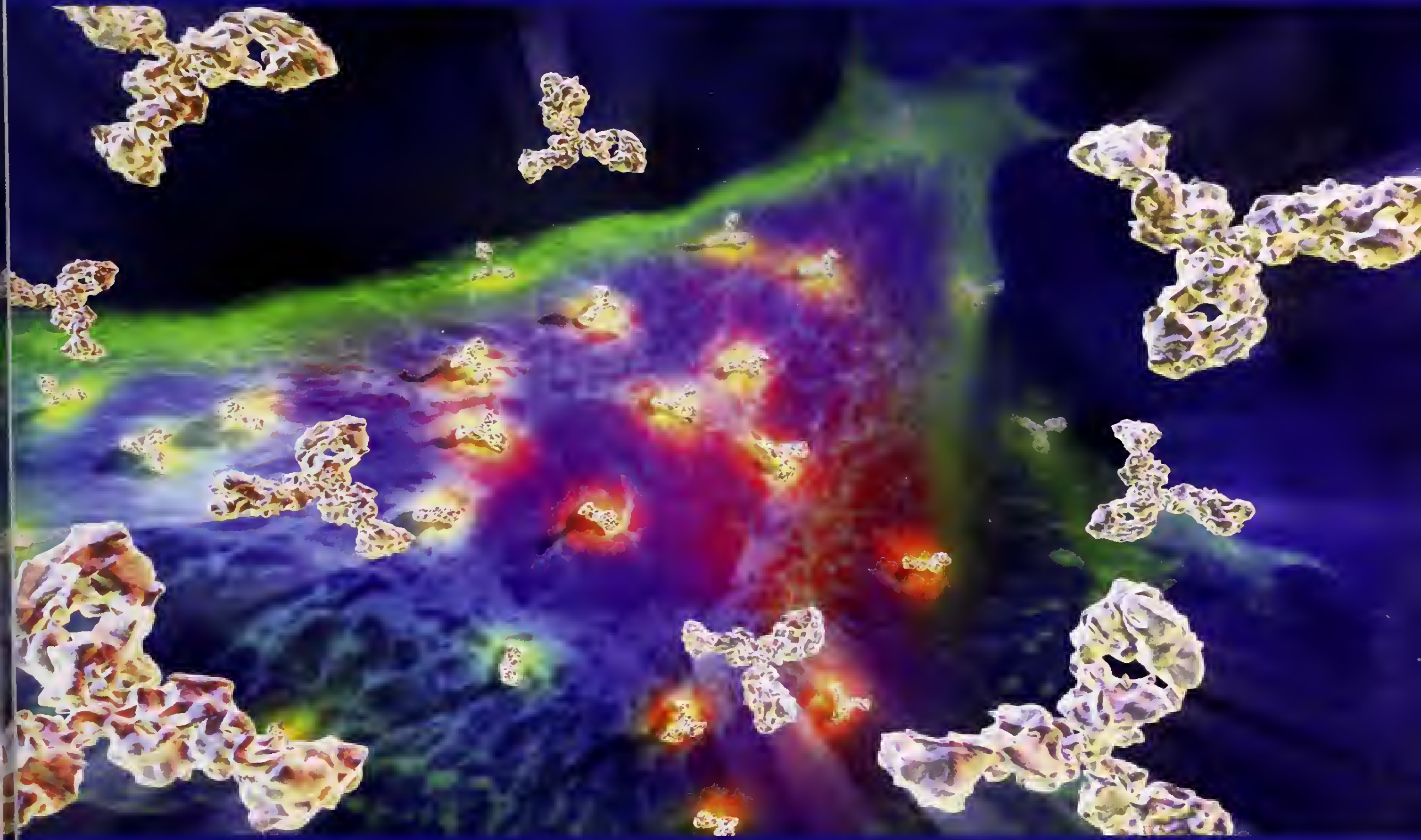
MOUNT SINAI JOURNAL OF MEDICINE

VOL. 77, NO. 6
NOVEMBER/DECEMBER 2010

A Journal of Translational and Personalized Medicine

Issue Theme
Cancer Therapy from Bench to Bedside

Theme Editors
William Oh, MD
Steven Burakoff, MD

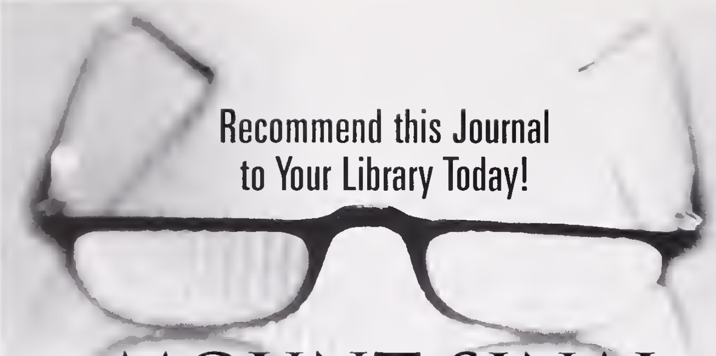


rgeted Inhibition of Kinases in Cancer Therapy
elanoma Immunotherapy
ostate Cancer Risk: Family History and Genetics
adder Cancer: Opportunities for Personalized Approach to Therapy
ng Adenocarcinoma: Lessons in Translation from Bench to Bedside
ew Translational Therapies–Pancreatic Cancer, Malignant Gliomas,
yeloproliferative Neoplasms, and Head and Neck Squamous Cell Carcinoma

pecial Features
Cross-cultural Perspective on Irritable Bowel Syndrome
scovering Neuroprotection in Parkinson's Disease, or Getting to Haphazard

WILEY-BLACKWELL

Editor-in-Chief: Penny A. Asbell, MD, FACS, MBA



Recommend this Journal to Your Library Today!

MOUNT SINAI JOURNAL OF MEDICINE

A Journal of Translational and Personalized Medicine

Editor-in-Chief: **Penny A. Asbell, MD, FACS, MBA**
Mount Sinai School of Medicine, New York, NY

Continuously published since 1934 and included in the *Index Medicus* from the very first volume, the **Mount Sinai Journal of Medicine (MSJM)** has a distinguished career as a general medical journal for a worldwide audience.

Dedicated to improving communication among all participants in medicine today, **MSJM** is an excellent source of up-to-date medical information for healthcare clinicians, educators, researchers and policy makers worldwide. **MSJM** will inspire the creative thinking needed to create new ways to understand, prevent, and treat human disease.

Each issue includes several quality, in-depth review articles related to a single theme and several others on subjects of diverse clinical interest. All theme articles are written by renowned authorities in their fields, from both Mount Sinai and elsewhere, under the supervision of Theme Editors.

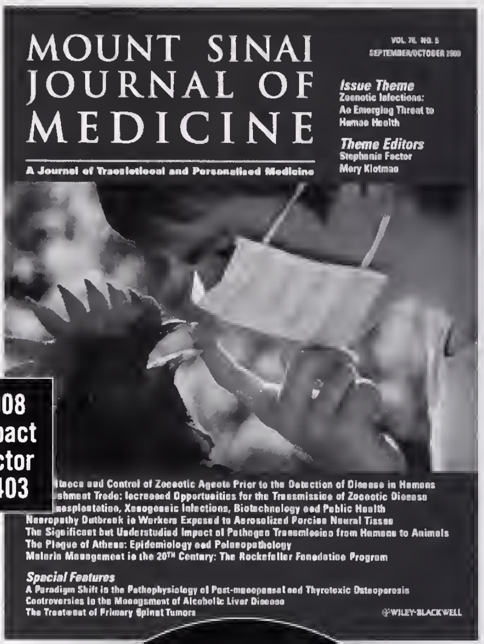
Upcoming Themes for Volume 77:

- Alzheimer's Disease: New Light on Etiology and Management
- Genes, Race, Ethnicity and the Environment
- Endovascular Surgery: Advances in Science and Technique
- Drug Discovery in Academic Settings
- The Obesity Pandemic
- Children's Health and the Environment

ACCESS THESE 2009 EDITORS CHOICE ARTICLES FREE ONLINE

- Advances and Controversies in Assisted Reproductive Technology (76:6)
- Xenotransplantation, Xenogenic Infection, Biotechnology and Public Health (76:5)
- The Student-run Health Clinic: A Novel Arena to Educate Medical Students on Systems-based Practice (76:4)
- Glycemic Targets for Patients with Type 2 Diabetes Mellitus (76:3)
- Neurobehavioral Sequelae of Traumatic Brain Injury (76:2)
- Spectrum of Fibrosing Diffuse Parenchymal Lung Disease (76:1)

Want more? Access a free sample issue online today!


MOUNT SINAI JOURNAL OF MEDICINE
A Journal of Translational and Personalized Medicine
VOL. 76, NO. 5
SEPTEMBER/OCTOBER 2009
Issue Theme
Zoonotic Infections: An Emerging Threat to Human Health
Theme Editors
Special Features
WILEY-BLACKWELL

2008 Impact Factor 1.403

Stay on Top of the Research!

Sign-up for Email Table of Contents Alerts delivered direct to your desktop – FREE!

Simply go to the homepage and click on "Set E-Mail Alert" for automatic table of contents alerts, or create a customized alert based on your own search criteria.

How to Receive Complimentary Online Access for Institutions?

Ask your librarian to e-mail optinaccess@wiley.co.uk or visit www.interscience.wiley.com/msjregister for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Library Recommendation Form

To______________________________

Date______________________________

Requested by:______________________________

Dept.______________________________

I have reviewed the **Mount Sinai Journal of Medicine** and urge you to add a complimentary 2010 online subscription to our subscription list. Please contact me should you have any questions.

WWW.MOUNTSINAIJOURNAL.ORG

WILEY-BLACKWELL

Enter Your 2010 Subscription Today!

MOUNT SINAI JOURNAL OF MEDICINE
A Journal of Translational and Personalized Medicine
Issue Theme
Type 2 Diabetes Mellitus:
New Evidence,
More Controversy
Theme Editor
Dina Green
Special Features
©WILEY-BLACKWELL

imed at clinicians and educators, the **ount Sinai Journal of Medicine** provides igh-quality, in-depth review articles in very issue, studded with clinical pearls.

our 2010 will include the six 2010 print ssues of the **Mount Sinai Journal of edicine** (Volume 77), plus concurrent ll-text online access to these six issues n Wiley's Internet publishing platform.

tandard Personal Subscriptions

130.00 (U.S./Canada/Mexico)
120.00 (Outside North America)

ount Sinai Alumni Subscriptions!

0.00 (Worldwide)

xtended for Institutions in 2010

omplimentary nline Access!

ur institution can opt to receive mplimentary online access to the **ount Sinai Journal of Medicine** during 10. When registering for complimentary line access, your institution can also tain one complimentary print bscription.

PERSONAL SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 personal subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $105.00 (U.S./Canada/Mexico) ☐ $130.00 (Outside North America)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

Method of Payment (Canada residents please add GST. Personal rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **John Wiley & Sons, Inc.** Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA

Card #____________________ Exp. Date__________

Signature______________________________
(credit card orders not valid without signature)

Ship To:

Name______________________________

Address______________________________

City____________________ State________ Zip__________

MSJM0910

MSJM ALUMNI SUBSCRIPTION ORDER FORM

☐ **Yes**, please enter my 2010 Alumna/Alumnus Subscription to the **Mount Sinai Journal of Medicine** at the following subscription rate.

☐ $40.00 (Worldwide)

Please note that all subscriptions are entered on a calendar year basis, regardless of the month that the subscription form is sent back. You will receive all issues from January 2010 through December 2010.

METHOD OF PAYMENT (Alumna/Alumnus rate subscriptions must be paid by personal check or credit card. They may not be resold or used as library copies.)

☐ Check/Money Order (Payable to **The Mount Sinai School of Medicine** Payment must be made in U.S. dollars and drawn on a U.S. bank.)

☐ Charge My Credit Card: ☐ American Express ☐ MasterCard ☐ VISA ☐ Other

Card #____________________ Exp. Date__________

Signature______________________________
(credit card orders not valid without signature)

Ship To:

Name______________________________

Address______________________________

City____________________ State________ Zip__________

To subscribe online visit: **www.msjmonline.com**

MSJM10

COMPLIMENTARY ONLINE ACCESS FOR INSTITUTIONS

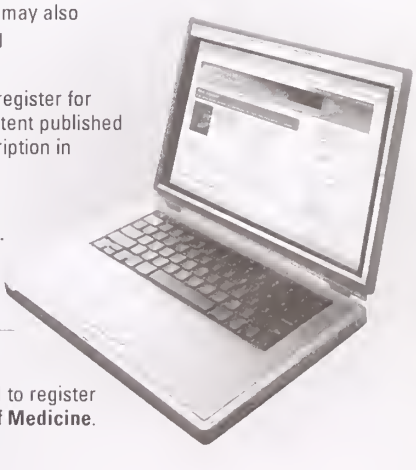
Mount Sinai Journal of Medicine

A Journal of Translational and Personalized Medicine

2010. Volume 77. 6 issues per year. Print ISSN: 0027-2507. Online ISSN: 1931-7581.

Your institution can opt to receive complimentary online access to the **Mount Sinai Journal of Medicine** during 2010. Your institution may also request one complimentary print subscription when registering for complimentary online access.

There is no obligation to subscribe in subsequent years if you register for complimentary access. Institutions retain online access to content published during the complimentary period, with or without a paid subscription in subsequent years.



Dear Librarian,
I have reviewed the scope of Mount Sinai Journal of Medicine.
Please register for complimentary access:

Name: _______________ Department. __________

Signature: _______________ Date: __________

Ask your librarian to e-mail **optinaccess@wiley.co.uk** or visit **www.interscience.wiley.com/msjregister** for more details and to register for complimentary online access to the **Mount Sinai Journal of Medicine**.

Don't delay, recommend to your library today!

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS MAIL PERMIT NO. 2277 HOBOKEN, NJ

POSTAGE WILL BE PAID BY ADDRESSEE

JOHN WILEY & SONS INC
ATTN: BETH MOLD, MAIL STOP 8-01
111 RIVER STREET
HOBOKEN, NJ 07030-9441

PLACE
POSTAGE
STAMP
HERE

INES SANTOS
MOUNT SINAI JOURNAL OF MEDICINE
ONE GUSTAV L. LEVY PLACE BOX 1094
NEW YORK, NY 10029-6574

Library Routing Card

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

Editor-in-Chief

Penny A. Asbell

Deputy Editor

David H. Adams

Associate Editors

Stuart Aaronson
Erwin Bottinger
Dennis S. Charney
Valentin Fuster
Wayne Goodman
Jonathan Halperin
Ravi Iyengar
Ihor R. Lemischka
Michael L. Marin
Thomas G. McGinn
Ira Nash
Maria New
Rosamond Rhodes
Hugh Sampson
Mary Sano
Jeffrey H. Silverstein
Albert Siu
Anthony Squire
Fred Suchy
Savio Woo

Editorial Board

William J. Bremner	University of Washington
Thomas M. Coffman	Duke University
Dwight Evans	University of Pennsylvania
Gabriel Haddad	University of California, San Diego
Heidi E. Hamm	Vanderbilt University
Matti Häyry	University of Manchester
Eve J. Higginbotham	Howard University
Katherine A. High	University of Pennsylvania
Tony Hope	University of Oxford
Mary Klotman	Duke University
Paul Klotman	Baylor College of Medicine
John H. Krystal	Yale University
Raju Kucherlapati	Harvard Medical School
Robert J. Lefkowitz	Duke University
Robert Lisak	Wayne State University
George Lister	University of Texas
Philip A. Marsden	University of Toronto
Charles Nemeroff	Emory University
Scott Rauch	Harvard Medical School
Bernard Roizman	University of Chicago
Alan F. Schatzberg	Stanford University
Inder Verma	The Salk Institute
Jeffrey A. Whitsett	University of Cincinnati
Allan Wolkoff	Albert Einstein College of Medicine
Noam Zohar	Bar-Ilan University

Theme Editors

William K. Oh, M.D.
Professor of Clinical Cancer Therapeutics
Associate Director, Tisch Cancer Institute

Steven J. Burakoff, MD
Professor of Oncological Sciences
Director, Tisch Cancer Institute

Mount Sinai School of Medicine

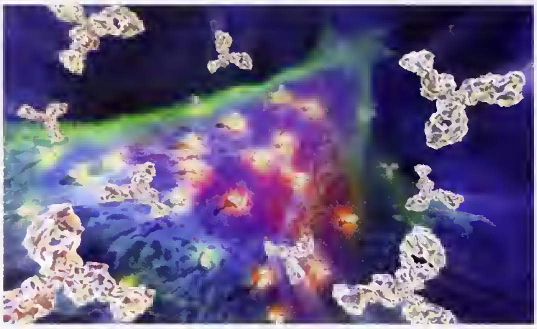
Editorial Director

Stephen Cohen
stephen.cohen@mssm.edu
212-241-6304

Publisher

John Wiley & Sons
Shawn Morton
Journal Publishing Director, Medicine
smorton@wiley.com

Cancer Therapy from Bench to Bedside



This artwork shows monoclonal antibodies approaching a cancer cell. The Y-shaped antibodies are engineered to be identical and specific to only one type of antigen on the cancer cell surface. Because of their specificity they can be used to deliver cytotoxic agents to cancer cells without affecting other cells in the body.

Future *MSJM* Themes and Theme Editors

- Future of primary care (Thomas McGinn)
- Patient safety (Ira Nash, Erin DePree)
- Children's environmental health (Philip Landrigan, Amir Miodovnik)
- Multiple Sclerosis (Fred Lublin, Tracey DeAngelis)
- Global health delivery and implementation (Rajesh Vedanthan)
- Atherosclerosis (Valentin Fuster, Zahi Fayad)
- Anesthesia and pain management (George Silvay, James Eisenkraft)

©2010 **Mount Sinai School of Medicine.** All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the publisher, or authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, Phone: (978) 750-8400; Fax: (978) 750-4770. Such requests and other permission inquiries should be addressed to the Permissions Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; Phone: (201) 748-6011; Fax: (201) 748-6008; E-mail: permreq@wiley.com. Printed in the United States of America by The Sheridan Press.

Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine (Print ISSN: 0027-2507; Online ISSN: 1931-7581 at Wiley Online Library, wileyonlinelibrary.com) is published bimonthly for the Mount Sinai School of Medicine by Wiley Subscription Services, a Wiley Company, 111 River Street, Hoboken, NJ 07030.

Send subscription inquiries to: Subscription Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-6645.

Offprint sales and inquiries should be addressed to the Reprint Billing Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-8776. All other inquiries should be addressed to the Customer Service Department at (201) 748-6645.

Individual subscription price: $95.00 in North America, $120.00 outside North America. Payment must be made in US dollars drawn on a US bank. Claims for undelivered copies will be accepted only after the following issue has been delivered. For more subscription prices and information, please visit www.mountsinaijournal.org.

Change of address: Please forward six weeks prior to move: present mailing label with change of address. Please enclose a copy of the mailing label or cite your subscriber reference number in order to expedite handling. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits. Send claims in care of: *Mount Sinai Journal of Medicine*, c/o John Wiley & Sons, Inc., Attn: Journals Admin Dept. UK, 111 River Street, Hoboken, NJ 07030.

Cancellations: Subscription cancellations will not be accepted after the first issue has been mailed. Periodicals postage paid at Hoboken, NJ and at additional mailing offices.

Postmaster: Send address changes to *Mount Sinai Journal of Medicine*, Subscription Distribution, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Printed in the United States of America. Single issues from current and recent volumes are available at the current single issue price from cs-journals@wiley.com. Earlier issues may be obtained from Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. Tel: +1 518 537 4700, Fax: +1 518 537 5899, Email: psc@periodicals.com

Advertising inquiries should be addressed to: Advertising Sales Manager, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Phone: (201) 748-6921; Fax: (201) 748-6207. Advertising in this journal is published with a caveat emptor understanding. The Publisher, Editors, and members of the Editorial Board do not imply warranty of products, nor quality, validity, or approval of education material offered by such advertising.

Indexed by: Biological Abstracts® (Thomson ISI), Chemical Abstracts Service/SciFinder (ACS), Current Contents®/Clinical Medicine (Thomson ISI), EMBASE/Excerpta Medica (Elsevier), Index Medicus/MEDLINE/PubMed (NLM), Science Citation Index Expanded™ (Thomson ISI), SCOPUS (Elsevier), Web of Science® (Thomson ISI).

This journal is printed on acid-free paper.

Disclaimer: All articles published, including but not limited to original research, reviews, clinical notes, clinical pearls, prismatic cases, clinician's corner, editorials, reports, and letters, represent the opinions and views of the authors and do not reflect any official policy or medical opinion of the Mount Sinai School of Medicine, the institutions with which the authors are affiliated, or the Publisher unless this is clearly specified. Articles published herein should not be relied upon as recommending or promoting a specific method, diagnosis, or treatment by physicians for any particular patient or disorder. Although the Editors and Publisher believe that drug selections and dosages and the specifications and usage of equipment and devices as set forth herein are in accord with current recommendations and practice at the time of publication, they accept no legal responsibility for any errors or omissions, and make no warranty, express or implied, with respect to material contained herein.

MOUNT SINAI JOURNAL OF MEDICINE

A JOURNAL OF TRANSLATIONAL AND PERSONALIZED MEDICINE

CONTENTS
VOL 77, NO 6

Theme Introduction

571 Cancer Therapy from Bench to Bedside
William Oh, Steven Burakoff

Theme Articles

573 Targeted Inhibition of Kinases in Cancer Therapy
Stacey J. Baker, Premkumar Reddy

587 Bladder Cancer: Current Management and Opportunities for a Personalized Approach
Matthew D. Galsky, Simon J. Hall

597 Lung Adenocarcinoma: Lessons in Translation from Bench to Bedside
Jaya Sangodkar, Suzanna Katz, Heather Melville, Goutham Narla

606 Advanced or Metastatic Pancreatic Cancer: Molecular Targeted Therapies
Soley Bayraktar, Caio M Rocha-Lima

620 Melanoma Immunotherapy
Shanthi Sivendran, Bradley Glodny, Michael Pan, Miriam Merad, Yvonne Saenger

643 Genetics of Prostate Cancer Risk
Mark M. Pomerantz, Matthew L. Freedman

655 Malignant Gliomas: New Translational Therapies
Joohee Sul, Howard A. Fine

667 Myeloproliferative Neoplasms: New Translational Therapies
John Mascarenhas, Ronald Hoffman

684 Head and Neck Squamous Cell Carcinoma: New Translational Therapies
Anthony Prince, Julio Aguirre-Ghizo, Eric Genden, Marshall Posner, Andrew Sikora

Special Features

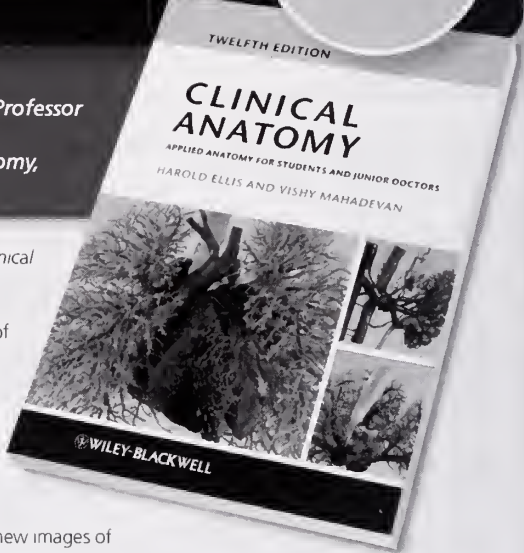
700 Discovering Neuroprotection in Parkinson's Disease, or Getting to Haphazard
Karl Kieburtz

707 A Cross-Cultural Perspective on Irritable Bowel Syndrome
Charles D. Gerson, Mary-Joan Gerson

Clinical Anatomy

Applied anatomy for students and junior doctors

TWELFTH EDITION


TWELFTH EDITION
CLINICAL ANATOMY
APPLIED ANATOMY FOR STUDENTS AND JUNIOR DOCTORS
HAROLD ELLIS AND VISHY MAHADEVAN
WILEY-BLACKWELL

Harold Ellis and Vishy Mahadevan

Guy's, King's and St Thomas' School of Biomedical Sciences; Emeritus Professor of Surgery, Charing Cross and Westminster Medical School, London; Barbers' Company Reader in Anatomy and Professor of Surgical Anatomy, The Royal College of Surgeons of England, London

Written by one of the great teachers of Anatomy, this 50th anniversary edition of *Clinical Anatomy* continues to provide thousands of medical students, postgraduate trainees and junior doctors across the world with essential anatomical information within a clinical setting. It is particularly appropriate for those preparing for the Membership of the Royal College of Surgeon's Exam.

For this new edition, Professor Harold Ellis is joined, for the first time, by a co-author. Professor Vishy Mahadevan is Barber's Company Reader in Anatomy at the Royal College of Surgeons, a renowned teacher of surgical trainees and is also an examiner for the MRCS Exam. Together they provide a detailed, easy-to-follow structured text suitable for anatomy students and trainees of all levels.

Fully revised and updated, this new edition features new and re-drawn illustrations, new images of normal anatomy (radiographs, CTs and NMRs) to aid understanding, a revised neuroanatomy section to provide greater clarity and an expansion of the lower limb section.

In order to provide revision support, *Clinical Anatomy* has its own website featuring digital flashcards of all the images found in the textbook available to download for free at **www.ellisclinicalanatomy.co.uk**

March 2010 • 9781405186179
470 pages

CONTENTS

Preface to the Twelfth (Jubilee) Edition
Preface to the First Edition
Acknowledgements to the Twelfth (Jubilee) Edition
Acknowledgements to the First Edition

Part 1: The Thorax
Introduction
Surface anatomy and surface markings
The thoracic cage
The lower respiratory tract
The mediastinum
On the examination of a chest radiograph

Part 2: The Abdomen and Pelvis
Surface anatomy and surface markings
The fasciae and muscles of the abdominal wall
Peritoneal cavity
The gastrointestinal tract
The gastrointestinal adnexae: liver, gall-bladder and its ducts, pancreas and spleen
The urinary tract
The male genital organs
The bony and ligamentous pelvis
The muscles of the pelvic floor and perineum
The female genital organs
The posterior abdominal wall
Computed axial tomography

Part 3: The Upper Limb
Surface anatomy and surface markings of the upper limb
The bones and joints of the upper limb
Three important zones of the upper limb: the axilla, the cubital fossa and the carpal tunnel
The arteries of the upper limb
The brachial plexus
The course and distribution of the principal nerves of the upper limb
Compartments of the upper limb
The female breast
The anatomy of upper limb deformities
The spaces of the hand

Part 4: The Lower Limb
The surface anatomy and surface markings of the lower limb
The bones and joints of the lower limb
Three important zones of the lower limb: the femoral triangle, adductor canal and popliteal fossa
The arteries of the lower limb
The veins of the lower limb
The course and distribution of the principal nerves of the lower limb
Compartments of the lower limb

Part 5: The Head and Neck
The surface anatomy of the neck
The thyroid gland
The palate
The tongue and floor of the mouth
The pharynx
The larynx
The salivary glands
The major arteries of the head and neck
The veins of the head and neck
The lymph nodes of the neck
The cervical sympathetic trunk
The branchial system and its derivatives
The surface anatomy and surface markings of the head
The scalp
The skull
The paranasal sinuses (accessory nasal sinuses)
The mandible
The vertebral column

Part 6: The Nervous System
The brain
The spinal cord
The cranial nerves
The special senses
The autonomic nervous system

Glossary of eponyms
Index

For more information or to place an order visit:
www.ellisclinicalanatomy.co.uk

WILEY-BLACKWELL

COMING SOON!
Download the FREE digital flashcards from *Clinical Anatomy*, at **www.ellisclinicalanatomy.co.uk**

THEME INTRODUCTION

Cancer Therapy from Bench to Bedside

William K. Oh, MD and Steven J. Burakoff, MD

Division of Hematology and Medical Oncology, Tisch Cancer Institute, Mount Sinai School of Medicine, New York, NY

We are in the midst of a revolution in clinical cancer therapeutics. For many decades, the only options for treating most cancers were radical surgery or radiation therapy; however, outcomes remained poor for the most common cancers, because many patients were diagnosed at more advanced stages of disease. Conventional cytotoxic chemotherapy has been associated with many advances, including significant improvements in the cure of testicular cancers, Hodgkin lymphoma, and acute lymphocytic leukemia in children. However, it has become clear that patients with many common diseases such as lung and pancreatic cancer have minimal benefit from traditional chemotherapy. It is in this light that a better understanding of the biology of cancer has helped to define better treatments for patients.

The most significant breakthrough in this regard came with the approval of imatinib mesylate (Gleevec) in 2001 for chronic myelogenous leukemia (CML), a disease that until then had been treated primarily with cytotoxic chemotherapy, with poor results. Imatinib was a rationally targeted small molecule inhibitor of the breakpoint cluster region-Abelson tyrosine kinase (bcr-abl), which was known to be the dominant genetic abnormality of CML, the so-called "Philadelphia chromosome." Even as patients experienced high response rates, however, resistance developed to imatinib and new drugs were required. Even here, though, understanding the molecular mechanisms of resistance to imatinib has been key to developing new kinase inhibitors such as dasatinib that translate into effective therapies for cancer patients in the modern era.

This, then, is the essence of translational medicine: the adaptation of new knowledge of the biology of cancer directly to patients who are battling cancer in the clinic every day. And despite the great breakthrough represented by Gleevec, it has remained challenging in the decade since to reproduce this success. Many cancers are more heterogeneous than CML, as might have been expected, and developing drugs against specific targets, even when the target is known, has been more difficult than some had hoped. Nonetheless, there recently has been an increasing sense of momentum in our understanding of the molecular heterogeneity of many common cancers, and the number of new drugs being rationally developed is increasing. For instance, breast cancer is no longer thought of as a single clinical entity, as it was only less than a decade ago. Patients with breast cancer are divided into different categories based on hormone-receptor status, Her-2/neu overexpression, and/or the presence of basal-like cells that confer relative hormone resistance. Today women are offered different treatment approaches based on their tumor subtype.

Some examples of these successes include targeting the epidermal growth factor receptor (EGFR) in lung and colon cancers, Her-2/neu in breast cancer, vascular endothelial growth factor and its receptor in renal cell carcinoma, and BRAF in malignant melanoma. Each of these targets has led to effective drugs with significant clinical activity and agents that are either approved or likely to be approved. The story of KRAS mutations is emblematic of the complex direction in which cancer therapeutics is headed. When mutated, the KRAS gene behaves as an oncogene (K-ras) and can induce multiple cancers, including colon and pancreatic cancer. The protein, K-RAS, is critical for EGFR signaling, so only

Address Correspondence to:

William K. Oh
Division of Hematology
and Medical Oncology
Mount Sinai School of Medicine
New York, NY
Email: william.oh@exchange.
mssm.edu

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20221

© 2010 Mount Sinai School of Medicine

patients who have wild-type K-RAS protein appear to respond to EGFR inhibitors such as cetuximab. Thus, intracellular signaling pathways in cancer cells determine susceptibility to targeted agents; for us to be as effective as possible, we must understand these interactions as we develop new drugs.

At the recent American Society of Clinical Oncology Annual Meeting in 2010, one of the key presentations involved the use of an ALK kinase inhibitor (crizotinib) in a population of lung cancer patients who harbor a fusion protein, EML4-ALK. Although only about 4% of all lung cancer patients in the United States are thought to have this molecular abnormality, >90% of those patients who do had stabilization or response of their cancers to treatment.

Another important area of targeted therapy will exploit the immune system to more specifically target cancer. In 2010, the US Food and Drug Administration approved the first cancer-specific immunotherapy treatment, sipuleucel-T (Provenge), for castration-resistant prostate cancer. Unfortunately, the mechanism of action for sipuleucel-T remains unclear, and further research will need to be done to better understand how it keeps men with this cancer alive longer, even as it does not induce tumor responses or delay progression of disease. Also, based on a recent large randomized trial, ipilimumab, an anti–cytotoxic T-lymphocyte antigen-4 antibody, will be approved for advanced metastatic melanoma. This immunotherapy is the first-ever treatment to show a survival benefit in metastatic melanoma and appears to work by enhancing T-cell responses to cancer. As we understand more about the specific immune responses to cancer, it may be possible for newer therapies to increasingly exploit the differences between cancer and normal cells by targeting the immune system.

Finally, the explosion of knowledge in genomic medicine is only slowly wending its way into clinical cancer care. Its impact can be felt in better molecular characterization of common cancers. As next-generation sequencing technologies improve the amount of knowledge we are able to acquire about the key aberrations in the cancer genome, soon the challenge will be in determining–in real time as patients are diagnosed–how that deluge of genetic information can be translated into a specific action plan that will optimize patient care.

In this issue of the *Mount Sinai Journal of Medicine*, our esteemed colleagues at the cutting edge of molecular medicine in cancer provide updates on many of these exciting therapeutic areas in cancer care for common solid tumors and hematological malignancies. We believe that this special theme issue provides a snapshot of a rapidly evolving field, one that we predict we will not recognize a decade from now because of the pace of new discoveries and new treatments. Our patients deserve no less than this expectation as we seek out better treatments for their cancers, based on an increasing fountain of knowledge of the underlying biology of cancer rather than a traditional approach of "one size fits all." This translational approach must continue at its current exponential growth if we are to make a major impact on the suffering and mortality from cancer in 2010 and beyond.

DOI:10.1002/MSJ

Targeted Inhibition of Kinases in Cancer Therapy

Stacey J. Baker and E. Premkumar Reddy

Mount Sinai School of Medicine, New York, NY

OUTLINE

BCR-ABL Oncogene–Targeted Therapy
  The Philadelphia Chromosome
  Imatinib
  Second-Generation BCR-ABL Inhibitors
  Overcoming T315I Resistance
Epidermal Growth Factor Receptor (EGFR)–Targeted Therapy
  EGFR and Cancer
  Gefitinib and Erlotinib: Small Molecule EGFR–Targeted Inhibitors
  Gefitinib and Erlotinib Resistance
  Overcoming T790M Resistance
Antibody-Based EGFR/HER-2/NEU/ERBB-2–Targeted Therapy
  HER-2 and Cancer
  Trastuzumab
  Other EGFR Family–Targeted Monoclonal Antibodies
Conclusion

ABSTRACT

With an understanding of the molecular changes that accompany cell transformation, cancer drug discovery has undergone a dramatic change in the past few years. Whereas most of the emphasis in the past has been placed on developing drugs that induce cell death based on mechanisms that do not discriminate between normal and tumor cells, recent strategies have emphasized targeting specific mechanisms that have gone awry in tumor cells. However, the identification of cancer-associated mutations in oncogenes and their amplification in tumors has suggested that inhibitors against such proteins might represent attractive substrates for targeted therapy. In the clinic, the success of imatinib (Gleevec®, STI571) and trastuzumab (Herceptin®), both firsts of their kind, spurred further development of new, second-generation drugs that target kinases in cancer. This review highlights a few important examples each of these types of therapies, along with some newer agents that are in various stages of development. Second-generation kinase inhibitors aimed at overriding emerging resistance to these therapies are also discussed. ***Mt Sinai J Med 77:573–586, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: kinase inhibitor, monoclonal antibody, small molecule inhibitors, targeted therapy.

Address Correspondence to:

E. Premkumar Reddy
Mount Sinai School of Medicine
New York, NY
Email: premkumar.reddy@mssm.edu

One of the most important breakthroughs in cancer research came from the discovery that oncogenes of acute transforming viruses are in fact derived form normal cellular DNA and that this genetic information is transduced by acute transforming viruses via genetic recombination. Studies within the last few decades have also shown that tumor cells often gain a growth advantage through amplification or mutation of various oncogenes, the end result of which appears to be a profound enhancement in the amplitude of the growth and survival signals. Some examples of these changes include amplification of the epidermal growth factor (EGF; also known as ErbB1) and ErbB2 receptors in lung and breast tumors, mutation of the *ras* oncogene in a wide variety of human tumors, and chromosomal translocations such as the Philadelphia chromosome (Ph), which results in the activation of the Abelson tyrosine kinase (Abl).

With an understanding of these molecular changes that accompany cell transformation, cancer drug discovery has undergone a dramatic change in the past few years. The elucidation of signaling pathways that are deregulated in tumor cells, as well as the identification of mutations in both oncogenes and growth-suppressor genes, has suggested multiple targets and revealed approaches for the development

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20220

© 2010 Mount Sinai School of Medicine

of new classes of drugs, from antibodies to receptors and small-molecule inhibitors to mutant kinases. The most successful of these types of agents, by far, is Gleevec® (imatinib mesylate, STI571; Novartis), and it is because of the tremendous success that this drug has had in the clinic that additional kinase inhibitors have been and are being developed. Because the past 5 years have witnessed a large amount of research performed in the area of rational drug design, it was not possible to review all of the approaches that are currently being developed. We have therefore limited this review to the discussion of a few rationally designed targeted therapies that have received approval by the US Food and Drug Administration and exemplify the utility and problems associated with this line of research.

BREAKPOINT CLUSTER REGION-ABL ONCOGENE–TARGETED THERAPY

The Philadelphia Chromosome

The Philadelphia chromosome (Ph) was discovered in 1960 by Nowell and Hungerford, who analyzed samples derived from 7 patients suffering from what was known at that time as chronic granulocytic leukemia.[1] Each patient harbored a similar "minute chromosome," and none showed any other chromosomal abnormality.[1] We now know that this abnormal Ph chromosome results from a reciprocal translocation between chromosomes 9 at band q34 and 22 at band q11. More importantly, this translocation fuses the breakpoint cluster region (*Bcr*) gene with the *Abl* gene and creates the BCR-ABL oncogene,[2] (Figure 1) whose expression is responsible for >90% of chronic myelogenous leukemias (CML).[3]

Imatinib

Until recently, CML was treated with a variety of chemotherapeutic and biotherapeutic drugs.[4] Because the BCR-ABL protein is active in the majority of CML cases, it has been possible to synthesize small molecules that inhibit BCR-ABL kinase activity in leukemic cells without adversely affecting the normal cell population. Imatinib (Figure 2) is a small molecule that binds to the kinase domain of BCR-ABL when the protein is in its closed, inactive conformation.[5] In this conformation, the

Imatinib is a small molecule that binds to the kinase domain of breakpoint cluster region (BCR)-ABL when the protein is in its closed, inactive conformation.

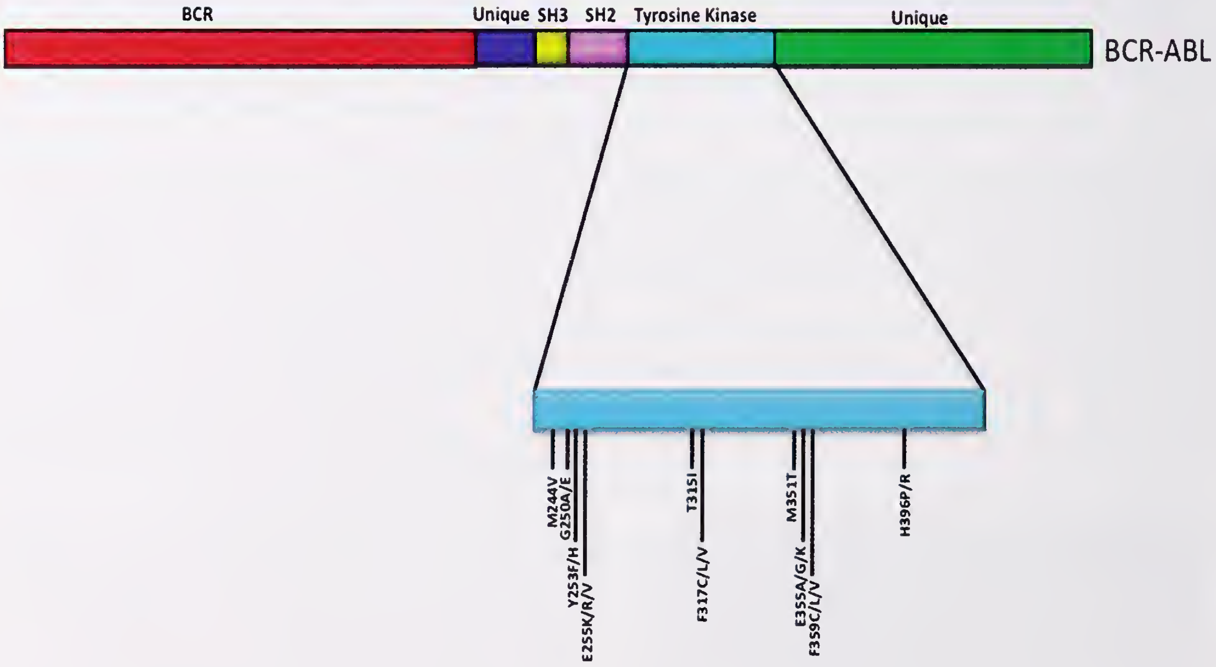


Fig 1. Schematic representation of the BCR-ABL protein. The positions of 10 of the most common mutations in the kinase domain that confer imatinib resistance are shown. (Note: Not drawn to scale). **Abbreviations:** BCR-ABL, breakpoint cluster region–Abelson tyrosine kinase.

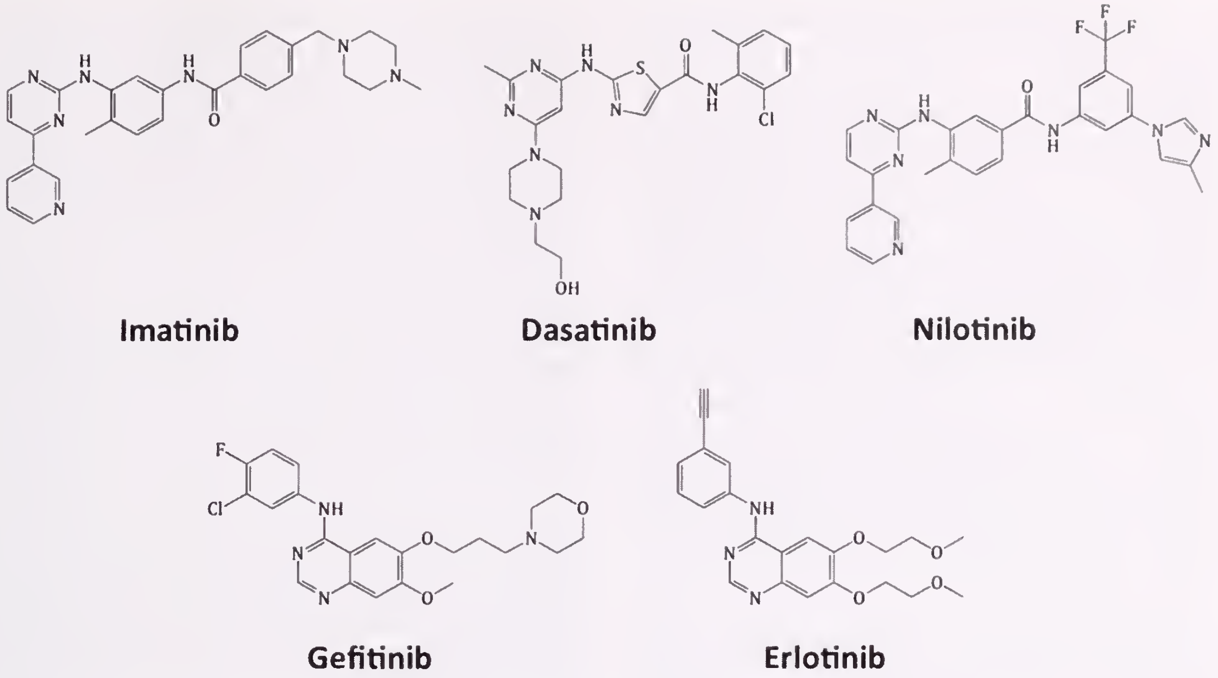


Fig 2. Structures of imatinib, nilotinib, dasatinib, gefitinib, and erlotinib.

catalytic central domain is blocked by the regulatory activation loop, and mutations within this loop have been shown to prevent the kinase from adopting an inactive conformation.[6]

As with most kinase inhibitors that are adenosine triphosphate (ATP) mimetics, imatinib inhibits several tyrosine kinases, including (but not limited to) platelet-derived growth factor receptor (PDGFR) α and β; c-Kit; Lck; fms; fibroblast growth factor receptor 1 (FGFR-1); vascular endothelial growth factor receptor (VEGFR) 1, 2, and 3; colony-stimulating factor-1 receptor (CSF-1R); and proto-oncogene c-RAF.[7,8] Imatinib also inhibits quinone oxidoreductase 2, even though it is not a kinase.[9] Imatinib, however, is most active against c-ABL, and more so its oncogenic forms. When BCR-ABL$^+$ cells are exposed to this drug, they do not proliferate and have been shown to undergo apoptosis, whereas normal, interleukin-3 (IL-3)-dependent cells remain virtually unaffected.[5,10]

In the clinic, the phase I trials aimed at assessing the safety of imatinib were remarkably successful. In those CML patients who were previously treated with interferon-alpha (IFN-α) but failed to respond, almost all of the patients who were treated with $\geq$300 mg imatinib achieved complete hematological responses. Furthermore, complete cytogenetic responses were observed in 13% of these patients. In phase II studies evaluating the efficacy of imatinib as a single agent, the results mirrored those obtained in phase I trials, whereby a significant number of patients in various stages of disease showed cytogenic responses. Phase III trials and those assessing the effectiveness of imatinib in combination with other cytotoxic agents and/or those aimed at improving the efficacy of imatinib monotherapy were equally successful, whereby imatinib proved to be the most effective treatment for various stages of leukemic disease. As a result, imatinib is now considered as a first-line therapy for the majority of CML cases due to its high efficacy level and relatively mild side effects.[7]

Imatinib Resistance in Chronic Myelogenous Leukemias

In spite of the fact that the majority of patients receiving imatinib respond to treatment at both the hematological and cytogenetic levels, relapse occurs in a subset of patients with chronic disease, and this number jumps significantly to nearly 100% in those patients whose disease is in the advanced stages.[3] Several studies have attempted to address the mechanism(s) by which CML cells acquire imatinib resistance.

Non–mutation-dependent mechanisms Elevated levels of BCR-ABL are sometimes expressed as a result of genomic amplification of the BCR-ABL locus[11–13] and amplification of the locus and/or multiple copies of the Ph chromosome have also reported in a small number of patients.[14,15] Studies have also revealed that imatinib resistance can be mediated by increased levels of the Pgp multidrug resistance protein (MDR1, ABCB1),[12] although patient-derived samples do not tend to support such a mechanism in a clinical setting.

Other studies have examined signaling pathways and proteins that are activated downstream of BCR-ABL. In some advanced stages of CML, levels of Hck and Lyn, members of the oncogenic protein kinase pp60^{src} (Src) family of tyrosine kinases, are increased and constitutively phosphorylated.[16–18] Additional BCR-ABL–independent mechanisms that mediate imatinib resistance include (but are not limited to) modulation of Janus kinase 2 (JAK2),[19] the activation of the Akt/phosphoinositide 3-kinase (PI3K)/mammalian target of rapamycin (mTOR) pathway,[20] activation of runt-related (Runx) transcription factors,[21] activation of mitogen-activated protein kinase (MAPK) activity,[22] and inactivation/mutation of p53.[23,24]

Mutation of BCR-ABL Despite these studies, the majority of reports indicate that the mechanism that accounts for the majority of imatinib-resistant leukemias in vivo is mutation of the BCR-ABL gene itself. Mutation within the kinase domain is most common (Figure 1). The first study that analyzed the frequency of these mutations was published in 2001 by Gorre *et al.*[14] and examined cells derived from 11 patients in the advanced stages of CML who were being treated with imatinib and had initially responded well to treatment but had relapsed. Although genomic amplification was detected in 3 of the samples, a specific mutation within the BCR-ABL kinase domain was present in 6 cases. Specifically, this mutation resulted in the substitution of a threonine residue with isoleucine at amino acid position 315 (T315I) of the kinase domain. This substitution, based on the crystal structure of the Abl kinase domain complexed with an imatinib derivative, is thought to decrease the binding affinity of imatinib to the BCR-ABL protein, whereby the isoleucine substitution prevents hydrogen bond formation between BCR-ABL and imatinib.[25]

> *The majority of reports indicate that the mechanism that accounts for the majority of imatinib-resistant leukemias in vivo is mutation of the BCR-ABL gene itself.*

Since that initial report, more than 50 mutations in the kinase domain of the BCR-ABL gene have been identified in imatinib-resistant leukemic cells. The list of mutations has been extensively summarized and reviewed in 2007 by Apperley.[6] Although the exact mechanism(s) that drive the mutation process are at present unclear, it is widely believed that many (if not all) result from the "selection pressure" exerted by exposure to imatinib. Studies indicate that they may also arise due to genetic instability induced by BCR-ABL[26,27] and inhibition of the mismatch and nucleotide excision repair processes.[28–30]

The imatinib-resistant mutations that arise in BCR-ABL can be classified as those that potentially interfere with the ability of imatinib to interact directly with the BCR-ABL kinase domain and those that are predicted to destroy or hinder the ability of the kinase domain to adopt a conformation that is required for imatinib binding.[3,31] The number of mutations that occur with the highest frequency tend to be clustered within the P-loop, the imatinib-binding site, the catalytic domain, and the activation loop. In patients, it is clear that certain mutations are present in particular phases of CML. However, it is at present unclear whether some or any of the mutations are causative of disease or whether they merely correlate with certain phases.[6]

> *The number of mutations that occur with the highest frequency tend to be clustered within the P-loop, the imatinib-binding site, the catalytic domain, and the activation loop.*

Second-Generation BCR-ABL Inhibitors

Because of the frequency of mutations, efforts have been focused on the identification of novel, second-generation BCR-ABL inhibitors that are active against imatinib-resistant mutants of the protein. The most successful alternatives to imatinib to date are second-generation inhibitors of BCR-ABL that inhibit the protein at lower concentrations than imatinib. Some agents, such as dasatinib, are unrelated to imatinib and are also inhibitory to other tyrosine kinases, making them active against a variety of imatinib-resistant tumors. Nilotinib, on the other hand, was developed as a rationally designed alternative to imatinib that exploited imatinib's affinity and selectivity for the Abl kinase domain.

Dasatinib

Originally termed BMS-354 825,[32,33] dasatinib (Sprycel®; Bristol-Myers Squibb) is an orally bioavailable inhibitor of BCR-ABL and was the first drug to be approved for the treatment of imatinib-resistant Ph$^+$ leukemias (Figure 2). The drug has a shorter half-life than imatinib and no active metabolites.[34] Although structurally unrelated to imatinib, dasatinib inhibits some of the same target proteins such as BCR-ABL (albeit at lower concentrations than imatinib), c-Kit, and PDGFR. It is also inhibitory to other kinases such as Src family tyrosine kinases, the Tec kinases Btk and Tec, and the ephrin A receptor tyrosine kinase.[9,32,35,36] Dasatinib does not share structural conformation requirements for BCR-ABL with imatinib and therefore binds to both active and inactive conformations of the protein.[37] As a result, the drug has been successful in inhibiting nearly all imatinib-resistant forms of the protein, with the exception of T315I.[33,38,39]

In the clinic, dasatinib has been successful in treating patients who were intolerant to imatinib or whose disease was resistant to imatinib therapy. In clinical trials, the drug has shown the greatest efficacy in patients in the chronic phase of CML. Nonetheless, cases of dasatinib resistance have begun to emerge. Specifically, sequencing of BCR-ABL$^+$ clones isolated from patients receiving dasatinib after failing or showing intolerance to imatinib revealed a predominance of mutations at the T315 and F317 residues. In addition to the T315I mutation, a second mutation at this residue, T315A, has also been detected in one patient.[40,41] Interestingly, this same mutation was isolated during a mutagenesis screen aimed at isolating dasatinib-resistant forms of BCR-ABL.[42] At the F317 residue, 4 different mutations have been identified (F317I/L/V/C), although the F317L mutation has been documented in the greatest percentage of cases. These studies suggest that these 2 residues may be key in the ability of dasatinib to inhibit BCR-ABL.[40,43] It is interesting to note that the F317L was tested as part of the panel of imatinib-resistant mutants in both *in vitro* cell-based assays as well as in animal tumor model systems.[33,38] Although it has been noted that higher amounts of dasatinib were required to inhibit kinase activity and cellular proliferation, the levels were not such that they substantially exceeded those for other imatinib-resistant mutations that do not arise in dasatinib-sensitive patients. Other studies have identified mutations at several additional residues,[40,44,45] and compound mutations have also been observed in patients undergoing sequential imatinib and dasatinib therapies. One of these mutations, V299L, is of interest as it occurs at a residue that directly binds to dasatinib. This mutation, however, is sensitive to both imatinib and nilotinib,[44] further underscoring the differences between these drugs.

Nilotinib

Nilotinib (Tasigna®, AMN107; Novartis)[46] is structurally related to imatinib, has a similar half-life, and has a greater affinity (approximately 20-fold) for wild-type BCR-ABL.[34,38] Because of the structural similarities between the 2 drugs (Figure 2), in addition to inhibiting BCR-ABL, nilotinib is also inhibitory to c-Kit, PDGFR, and the TEL-PDGFRα and FIP1L1-PDGFRα fusion proteins.[45,47] Also similar to imatinib, nilotinib binds to BCR-ABL when it is in an inactive conformation; however, its design allows it to bind more tightly, thereby enhancing its inhibitory activity.[48] As a result, most of the imatinib-resistant forms of BCR-ABL tested in *in vitro* cell-based assays are also 20-fold more sensitive to this drug, although studies using certain imatinib-sensitive cell lines have shown that the increased potency may be as high as 60-fold.[46,49] Nilotinib does, however, when compared with wild-type BCR-ABL, show a reduced ability to inhibit P-loop mutations, and this is correlated with a response to the drug in the clinic.[50] The only exception is the T315I mutation, which is completely insensitive to nilotinib.[46,51]

Overcoming T315I Resistance

The T315 residue is commonly referred to as the "gatekeeper" because the T315I mutation confers complete resistance to imatinib, dasatinib, and nilotinib. It has become clear that the use of ATP mimetics that have the ability to target the kinase domain drives the emergence of such mutations, and hence, compounds that inhibit the T315I conformation directly are invaluable. Use of x-ray crystallography in conjunction with structure-guided optimization has led to the isolation of SGX393, a compound that inhibits both T315I and T315A. It does not, however, inhibit certain other P-loop mutations. Fortunately, some of these mutations are inhibited by either nilotinib and/or dasatinib, suggesting that this agent, should it be used in a clinical setting, could be used in combination therapy.[52] Other drugs with the ability to inhibit T315I directly that are undergoing clinical trials are inhibitors of the Aurora kinases; these include MK-0457, XL228, PHA-739358, and KW-2449.[53,54] In addition to T315I inhibitors, it may also be possible to target signaling pathways that are either downstream or parallel to BCR-ABL. There are numerous compounds at various stages of development that could possibly fill such a need, including (but not limited to) Src family kinase inhibitors, inhibitors of HSP90, inducers of apoptosis, activators of protein phosphatases, and histone deacetylase inhibitors[53,54] and even an allosteric BCR-ABL inhibitor that binds to the kinase at its myristoylation site.[55]

EPIDERMAL GROWTH FACTOR RECEPTOR (EGFR)–TARGETED THERAPY

Epidermal Growth Factor Receptor (EGFR) and Cancer

The significance of the epidermal growth factor receptor (EGFR) in cancer was suggested years before it was shown that the protein was altered or aberrantly expressed in tumors. During the early 1980s, studies of the avian erythroblastosis virus led to the identification of viral, transforming versions of the EGFR,[56–58] and subsequent studies showed that introduction of the EGFR gene into NIH-3T3 cells results in ligand-dependent transformation.[59,60] In certain types of tumors, such as gliomas, the EGFR locus is amplified, resulting in high levels of the receptor. Other tumor types express mutant, constitutively active forms of the protein.[61] Activation of the EGFR stimulates proliferation, metastasis, angiogenesis, and other phenotypes that are beneficial for tumor growth. Epidermal growth factor receptor is expressed in a variety of cell types and tumors, including (but not limited to) those of the bladder, lung, head and neck, breast, colon, and pancreas. In certain types of malignancies, such as head and neck cancers, the level of EGFR expression is of prognostic value where elevated levels are correlated with decreased survival; whereas in others, this link is not as well defined.[61]

Gefitinib and Erlotinib: Small Molecule EGFR–Targeted Inhibitors

Gefitinib (ZD1839, Iressa®; AstraZeneca) (Figure 2) belongs to a class of compounds termed anilinoquinazolinones and is an inhibitor of EGFR enzymatic activity in the low nanomolar range.[62] Subsequent studies also revealed that gefitinib effectively inhibited the proliferation of a variety of established tumor cell lines,[63] although this did not necessarily correlate with EGFR expression.

Gefitinib, as well as most targeted kinase inhibitors developed to date, acts by reversibly competing with ATP for binding to the EGFR, thereby inhibiting its autophosphorylation and activity.[64] Cells that have been exposed to this drug undergo cell-cycle arrest and apoptosis[65,66] and express lower levels of proangiogenic proteins such as VEGF.[67,68] Epidermal growth factor has been shown to regulate the expression of VEGF, IL-8, and other angiogenic factors. It is therefore not surprising that gefitinib has the ability to regulate tumor growth by controlling the neovasculature.[69]

> *Gefitinib, as well as most targeted kinase inhibitors developed to date, acts by reversibly competing with adenosine triphosphate for binding to the epithelial growth factor receptor, thereby inhibiting its autophosphorylation and activity.*

Response to Gefitinib Treatment

In the United States, gefitinib was initially approved for the treatment of non-small-cell lung cancer (NSCLC) in patients who had failed or were not able to receive standard chemotherapeutic regimens. In general, nonsmokers, women, and Asians have an overall more favorable response to EGFR kinase inhibition.[70] However, overall response to gefitinib has been shown to vary considerably in these and other patient cohorts, thereby prompting research to identify biomarkers that could account for this variability and predict sensitivity. Sequence analysis of the EGFR in primary NSCLC tumors derived from patients in Japan and the United States[71,72] identified mutations within the tyrosine kinase domain that correlated with sensitivity to gefitinib. For the most part, the abnormalities are in-frame deletions in amino acids 747–750 (del746-750) or a point mutation at position 858 (L858R) Figure 3.[71–73] These studies also showed that such mutations, which result in increased levels of ligand-induced receptor autophosphorylation, are more common in nonsmokers, women, Japanese patients, and those with the adenocarcinoma form of the disease. Since these initial studies were performed, other mutations in the forms of in-frame deletions and insertions, point

> *Sequence analysis of the epidermal growth factor receptor in primary non-small-cell lung cancer tumors derived from patients in Japan and the United States identified mutations within the tyrosine kinase domain that correlated with sensitivity to gefitinib.*

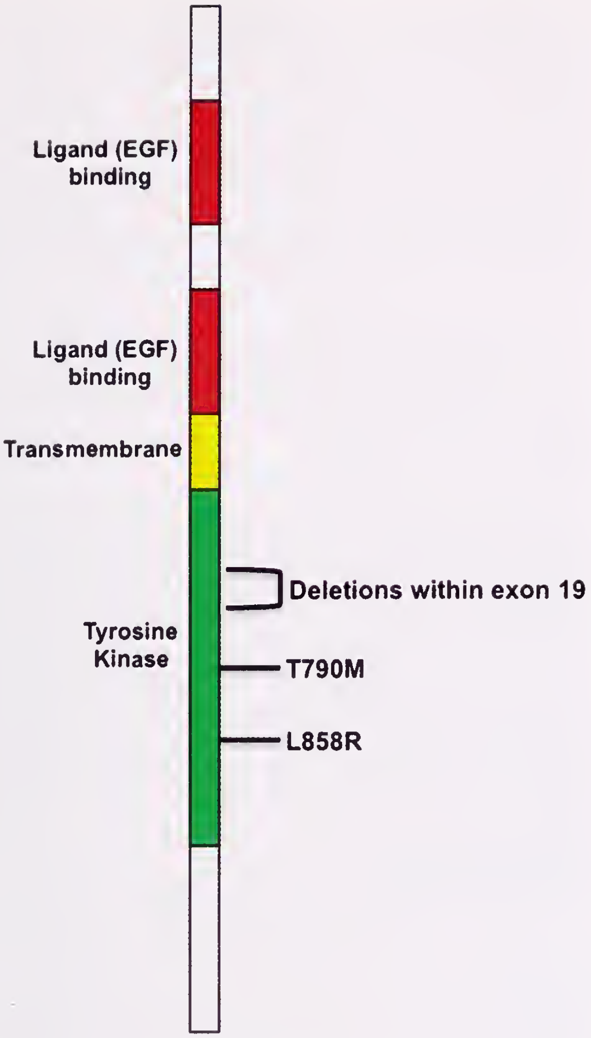


Fig 3. Schematic representation of the EGFR. The positions of the most commonly identified mutations are shown. (Note: Not drawn to scale). **Abbreviations:** EGF, epidermal growth factor; EGFR, epidermal growth factor receptor.

mutations, and duplications within exons 19–22 of the EGFR gene that confer some level of response to gefitinib have been identified. However, the "classical" L858R (exon 21) and the del746-750 (exon 19) mutations still account for roughly 90% of the alterations in this region.[74] Overall, NSCLC patients treated with gefitinib (or erlotinib) who harbor the del746-750 mutation have a longer median survival when compared with those with the L858R mutation.[75,76] In addition, immunohistochemical staining of tumor specimens for proteins that are downstream of the EGFR revealed that patients with tumors that stained positive for phosphorylated Akt (p-Akt) responded better to gefitinib treatment than those who stained negative; however, there was no such correlation for phosphorylated MAP kinase.[77] Similar results were obtained using human NSCLC cell lines, although these studies also revealed an association between p-Akt and ERK1/2 phosphorylation in gefitinib-resistant cells.[78,79]

Erlotinib

Erlotinib (Tarceva®, OSI-774, Genentech) (Figure 2) belongs to the class of quinazolinones and is a selective but reversible inhibitor of EGFR enzymatic activity in the low nanomolar range.[80] It also inhibits the constitutively active EGFRvIII variant that contains an in-frame deletion in exons 2–77 of the EGFR gene.[81] Erlotinib is currently approved for the treatment of advanced or metastatic resistant NSCLC patients and for use in combination therapy with gemcitabine in treating advanced, unresectable, or metastatic pancreatic cancer.[82,83] In *in vitro* assays, although the degree to which wild-type EGFR and the delL474-S752 mutant were inhibited by both gefitinib and erlotinib was similar, the L858R exhibited an approximate 10-fold increase in sensitivity to both drugs. It is therefore not surprising that in the clinic, patients whose tumors harbor mutant forms of the EGFR respond to treatment with erlotinib, which is similar to what has been observed with gefitinib.[73,83]

> *Erlotinib belongs to the class of quinazolinones and is a selective but reversible inhibitor of epidermal growth factor receptor enzymatic activity in the low nanomolar range.*

Gefitinib and Erlotinib Resistance

Secondary EGFR Mutations

Despite the success that these drugs have had in treating NSCLCs with EGFR-activating mutations, resistance emerges in a percentage of patients who have initially responded to treatment. Biochemical analysis and structural modeling have resulted in the identification of a second point mutation at amino acid position T790 (T790M) in a patient who had initially responded to gefitinib treatment but who had relapsed.[84] At that time, resistance was also predicted to occur in patients treated with erlotinib,[84] and, as expected, cases of erlotinib-resistant disease that harbor this mutation have emerged in the clinic.[85] The same mutation has also been identified in a family with several cases of inherited NSCLC and is thought to act *in cis* with a second somatic T790M mutation.[86]

In cis cooperation has also been observed with both the L858R and del746-750 EGFR mutations as is evidenced by increased transformation potential and ligand-independent signaling.[87] Termed as EGFR's "gatekeeper," T790 is analogous to T315 of BCR-ABL as it is necessary to maintain proper conformation of the ATP-binding cleft and because of its ability to confer resistance to gefitinib and erlotinib. However, T790M does differ from T315I in that it does not have increased enzymatic activity *in vitro*.[85] At a basic mechanistic level, T790M increases the affinity of the EGFR kinase domain for ATP[88] and this single mutation accounts for approximately 50% of all cases of resistance to both gefitinib and erlotinib.[89,90] In addition to T790M, other mutations that arise in TKI-resistant tumors inhibition have been isolated in rare instances.[90]

Non-EGFR Mutation Dependent Mechanisms

In addition to mutation of the EGFR, mutations in *RAS* family genes are present in approximately 15%–30% of lung adenocarcinomas in patients with a history of cigarette smoking. Most of these mutations are located in exons 12 and 13 of the KRAS gene and are associated with resistance to gefitinib and erlotinib.[91]

Activation of parallel signaling pathways or those that are downstream of the EGFR can also result in resistance to gefitinib and erlotinib. *MET* locus amplification was originally identified in a gefitinib-resistant cell line and the resultant increased levels of MET protein signal through Akt via ErbB3. Subsequent analysis revealed that the same pathway was activated in the majority of gefitinib and erlotinib resistant tumors with *MET* amplification and that this amplification occurs in the presence or absence of the T790M mutation.[92,93] Other proteins/pathways whose activation is associated with EGFR kinase inhibition include (but are not limited) to PI3K/Akt,[94] the insulin-like growth factor 1 receptor (IGF-1R),[95,96] s-Src (in conjunction with c-Met),[97] and MAPK.[95]

Overcoming T790M Resistance

Overcoming resistance to the T790M mutation is of critical importance in the area of EGFR kinase inhibition. This mutation is a "generic" mutation and biochemical assays and x-ray crystallization predict that it will confer resistance or at least reduce the inhibitory properties of most if not all ATP-competitive inhibitors.[88] Both erlotinib and gefitinib are reversible inhibitors of the EGFR, and T790M reduces the affinity of the EGFR for both. Second-generation irreversible EGFR inhibitors, such as HKI-272,[98,99] have shown efficacy in mouse models and are predicted to overcome this resistance through their ability to form covalent bonds with the EGFR protein.[88,100] Although these results are encouraging, resistance to HKI-272 recently has been shown in in vitro cell culture systems,[101] and it is predicted that other irreversible EGFR inhibitors will be required to overcome the effects of T790M.

ANTIBODY-BASED EGFR/HER-2/NEU/ERBB-2–TARGETED THERAPY

HER-2 and Cancer

Heregulin-2 (HER-2), also known as ErbB-2 and Neu, is a member of the EGFR family of receptor tyrosine kinases. 102,103 ErbB-2 was first shown to be an oncogene in 1987 using NIH-3T3 transformation assays.[60,104] Soon after, mouse mammary tumor virus–driven HER-2 expression was shown to cause mammary tumors in transgenic mice,[105] providing early experimental evidence that suggested a role for this gene in the genesis and progression of breast cancer.

Trastuzumab

In human cancers, particularly those of the breast, amplification of HER-2 results in its overexpression (approximately 25%–30%) and is correlated with poor prognosis, tumor grade, likelihood of relapse and rate of survival.[106–108] High levels of HER-2 expression have also been observed in other tumor types and like those of the breast, such expression is also correlated with poorer outcomes.[109] It is believed that this overexpression results in constitutive activation via the formation of HER-2 homodimers.[110] As a result, many groups focused their efforts on the development of an antibody that could bind to the extracellular domain of HER-2 and block the adverse effects of its overexpression. Several murine monoclonal antibodies were capable

> *The humanized, chimeric version of this antibody, termed trastuzumab actually has a higher affinity for HER-2 than its parent and shares its ability to suppress the proliferation of HER-2$^+$ cells and xenografts as opposed to cells that do not overexpress the receptor.*

of inhibiting the proliferation of cell lines and human breast cancer xenografts that overexpressed the receptor, but the most promising candidate in this regard proved to be Genentech's MAb 4D5.[111] The humanized, chimeric version of this antibody, termed trastuzumab, or Herceptin® (Genentech), actually has a higher affinity for HER-2 than its parent and shares its ability to suppress the proliferation of HER-2^+ cells and xenografts as opposed to cells that do not overexpress the receptor.[111–113]

Although the exact mechanism by which trastuzumab acts is not completely defined, studies have shown that exposure to the antibody, which binds to the extracellular domain of HER-2, inhibits dimerzation of the receptor and thereby attenuates HER-2 signaling.[114,115] As a result, the proteins that are activated downstream of HER-2 that promote proliferation, such as PI3K, are effectively inhibited as well.[116] Those cells that have ceased proliferating become arrested in the G_1 phase of the cell cycle via increased levels of $p27^{Kip1}$/cdk-2 complexes,[116–120] effects that are enhanced by extending the exposure time to the antibody and increasing the dosage.[121] Trastuzumab also antagonizes tumor growth by inhibiting the expression of proangiogenic factors such as VEGF[122] and inducing others that are antiangiogenic such as thrombospondin 1.[123] Finally, proteolytic cleavage of the extracellular domain of HER-2 is also inhibited by trastuzumab.[124]

Resistance

In the clinic, when used as a single agent, trastuzumab is effectively used in metastatic breast cancer patients who have undergone previous treatment regimens with as well as those who were untreated, resulting in overall responses of 20%–35% and 11%–18%, respectively. These percentages improve when the drug is used in combination therapy.[125–127] Based on this and other data obtained in clinical trials, patients with HER-2^+ early-stage breast cancer are placed on a 1-year course of trastuzumab therapy. Combination therapy is also indicated for those patients with HER-2^+ metastatic disease. Despite these response rates, resistance to trastuzumab emerges in a percentage of patients who have been treated with the drug within a period of 1 year.[109] The emergence of resistance over the course of treatment, taken together with the fact that some tumors are inherently insensitive to trastuzumab, underscores the need to understand the pathways that are associated with a positive response to the drug.

At a mechanistic level, some of the first *in vitro* studies with breast tumor cell lines revealed that increased levels of IGF-1R antagonize the effects of trastuzumab.[128] These cells, which overexpressed the IGF-1R gene, also showed reduced levels of $p27^{Kip1}$, a phenotype that was also observed in resistant forms of the SKBR3 cell line that overexpress HER-2.[129] In these cells, IGF-1R leads to the up-regulation of the Skp2 ubiquitin ligase that mediates degradation of p27.[129–131]

Another protein that has been shown to mediate a clinical response to trastuzumab is phosphatase and tensin homolog (PTEN), a lipid phosphatase that inhibits proliferation by inhibiting PI3K-mediating signaling.[132] Studies by Nagata *et al.*[133] have shown that treatment with trastuzumab reduces PTEN phosphorylation and increases its membrane localization and activity within 10 minutes of administration. Patients with breast cancers that do not express PTEN had poorer responses to treatment than those with normal PTEN levels. Furthermore, PTEN-deficient tumors regained their responsiveness to trastuzumab when treated with PI3K inhibitors.[133] It was later demonstrated that trastuzumab inhibits the ErbB-3/PI3K/Akt pathway *in vitro* in a variety of tumor types and that tumors that harbor activating mutations in PI3K (PI3KCA) are insensitive to the drug. As PI3K is normally down-regulated in response to trastuzumab, these results indicate that trastuzumab resistance can be overcome with coadministering such inhibitors.[114]

Resistance to trastuzumab is also observed tumors that express variants of HER-2 that lack the trastuzumab-binding site in the extracellular domain, as well as high levels of transforming growth factor-α (TGF-α) and MUC-4, both of which inhibit binding of trastuzumab to the receptor.[109] In cases such as these where trastuzumab simply cannot bind to the receptor, use of second-generation EGFR/ErbB-2 small molecule inhibitors, such as lapatinib (GlaxoSmithKline), may be of particular interest.

Other EGFR Family–Targeted Monoclonal Antibodies

Cetuximab (IMC, Erbitux®; Bristol-Myers Squibb, Merck, Imclone Systems) is a chimeric monoclonal IgG_1 antibody that was approved in 2004 as either monotherapy or as part of a combination-therapy regimen with irinotecan for metastatic colorectal cancer patients and induces receptor internalization and antibody-based cytotoxicity.[134] Because subsequent clinical trials showed that the drug was efficacious against other solid tumors, cetuximab was later approved for use in combination therapy with

DOI:10.1002/MSJ

radiation or as a single agent in patients who previously underwent platinum-based chemotherapy.[135] Interestingly, unlike many EGFR-targeted antibodies, cetuximab also has demonstrable activity against the mutant vIII variant of the receptor that is expressed in NSCLC, inhibiting its kinase activity and internalization.[136–138] This property suggests that cetuximab could be useful in a greater variety of tumor types, including those that express mutant L858R/T790M EGFRs, that are resistant to small molecule kinase inhibitors.[136]

Another anti-EGFR monoclonal antibody currently approved for the treatment of metastatic colorectal cancer is panitumumab (ABX-EGF-Vectibix®; Amgen), a human IgG_2 monclonal antibody that binds to the extracellular domain of the EGFR and inhibits its activity and internalization. Panitumumab was the first recombinant human antibody approved for the treatment of $EGFR^+$ metastatic colorectal cancers that were refractory to previous chemotherapy regimens that included either 5-fluorouracil, irinotecan, and oxaliplatin.[139] Unfortunately, patients with mutations in KRAS do not benefit from cetuximab or panitumumab treatment,[140] underscoring the need to use these agents in patients with wild-type KRAS or to use other small molecule inhibitors that suppress Ras-mediated signaling pathways.

CONCLUSION

Targeted therapies represent some of the best discoveries that have resulted from the convergence of basic and clinical research efforts. The development of imatinib and trastuzumab, both firsts of their kind, have revolutionized the treatment of cancer and spawned the development of second-generation inhibitors. Though it is clear that we are witnessing a change in the paradigm of cancer therapy, resistance to these inhibitors presents a challenge to researchers and physicians alike. The emergence mutations in the ATP-binding domains of tyrosine kinases and the activation of downstream and parallel signaling pathways in drug-resistant tumors highlight the difficulty in generating inhibitors that can effectively overcome resistant forms of the disease and remind us that there is, unfortunately, always a need to develop novel targeted therapies.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health (1R01CA109820-04 and 5 R01 HL080666-03) to E. P. R.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Nowell PC, Hungerford DA. A minute chromosome in human granulocytic leukemia. *Science* 1960; 132: 1497.
2. Heisterkamp N, Stam K, Groffen J, et al. Structural organization of the bcr gene and its role in the Ph′ translocation. *Nature* 1985; 315: 758–761.
3. Shah NP, Sawyers CL. Mechanisms of resistance to STI571 in Philadelphia chromosome–associated leukemias. *Oncogene* 2003; 22: 7389–7395.
4. Hehlmann R. Current CML therapy: progress and dilemma. *Leukemia* 2003; 17: 1010–1012.
5. Druker BJ, Tamura S, Buchdunger E, et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. *Nat Med* 1996; 2: 561–566.
6. Apperley JF. Part II: management of resistance to imatinib in chronic myeloid leukaemia. *Lancet Oncol* 2007; 8: 1116–1128.
7. Deininger M, Buchdunger E, Druker BJ. The development of imatinib as a therapeutic agent for chronic myeloid leukemia. *Blood* 2005; 105: 2640–2653.
8. Mashkani B, Griffith R, Ashman LK. Colony-stimulating factor-1 receptor as a target for small molecule inhibitors. *Bioorg Med Chem* 2010; 18: 1789–1797.
9. Hantschel O, Rix U, Schmidt U, et al. The Btk tyrosine kinase is a major target of the Bcr-Abl inhibitor dasatinib. *Proc Natl Acad Sci U S A* 2007; 104: 13283–13288.
10. Deininger MW, Goldman JM, Lydon N, et al. The tyrosine kinase inhibitor CGP57148B selectively inhibits the growth of BCR-ABL-positive cells. *Blood* 1997; 90: 3691–3698.
11. Le Coutre P, Tassi E, Varella-Garcia M, et al. Induction of resistance to the Abelson inhibitor STI571 in human leukemic cells through gene amplification. *Blood* 2000; 95: 1758–1766.
12. Mahon FX, Deininger MW, Schultheis B, et al. Selection and characterization of BCR-ABL positive cell lines with differential sensitivity to the tyrosine kinase inhibitor STI571: diverse mechanisms of resistance. *Blood* 2000; 96: 1070–1079.
13. Weisberg E, Griffin JD. Mechanism of resistance to the ABL tyrosine kinase inhibitor STI571 in BCR/ABL-transformed hematopoietic cell lines. *Blood* 2000; 95: 3498–3505.
14. Gorre ME, Mohamed M, Ellwood K, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* 2001; 293: 876–880.
15. Hochhaus A, Kreil S, Corbin AS, et al. Molecular and chromosomal mechanisms of resistance to imatinib (STI571) therapy. *Leukemia* 2002; 16: 2190–2196.
16. Donato NJ, Wu JY, Stapley J, et al. BCR-ABL independence and LYN kinase overexpression in chronic myelogenous leukemia cells selected for resistance to STI571. *Blood* 2003; 101: 690–698.

17. Donato NJ, Wu JY, Stapley J, et al. Imatinib mesylate resistance through BCR-ABL independence in chronic myelogenous leukemia. *Cancer Res* 2004; 64: 672–677.
18. Wu J, Meng F, Kong LY, et al. Association between imatinib-resistant BCR-ABL mutation-negative leukemia and persistent activation of LYN kinase. *J Natl Cancer Inst* 2008; 100: 926–939.
19. Wang Y, Cai D, Brendel C, et al. Adaptive secretion of granulocyte-macrophage colony-stimulating factor (GM-CSF) mediates imatinib and nilotinib resistance in BCR/ABL$^+$ progenitors via JAK-2/STAT-5 pathway activation. *Blood* 2007; 109: 2147–2155.
20. Burchert A, Wang Y, Cai D, et al. Compensatory PI3-kinase/Akt/mTor activation regulates imatinib resistance development. *Leukemia* 2005; 19: 1774–1782.
21. Miething C, Grundler R, Mugler C, et al. Retroviral insertional mutagenesis identifies RUNX genes involved in chronic myeloid leukemia disease persistence under imatinib treatment. *Proc Natl Acad Sci U S A* 2007; 104: 4594–4599.
22. Chu S, Holtz M, Gupta M, et al. BCR/ABL kinase inhibition by imatinib mesylate enhances MAP kinase activity in chronic myelogenous leukemia CD34$^+$ cells. *Blood* 2004; 103: 3167–3174.
23. Wendel HG, de Stanchina E, Cepero E, et al. Loss of p53 impedes the antileukemic response to BCR-ABL inhibition. *Proc Natl Acad Sci U S A* 2006; 103: 7444–7449.
24. Burchert A. Roots of imatinib resistance: a question of self-renewal? *Drug Resist Updat* 2007; 10: 152–161.
25. Schindler T, Bornmann W, Pellicena P, et al. Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science* 2000; 289: 1938–1942.
26. Jiang X, Saw KM, Eaves A, et al. Instability of BCR-ABL gene in primary and cultured chronic myeloid leukemia stem cells. *J Natl Cancer Inst* 2007; 99: 680–693.
27. Jiang X, Zhao Y, Smith C, et al. Chronic myeloid leukemia stem cells possess multiple unique features of resistance to BCR-ABL targeted therapies. *Leukemia* 2007; 21: 926–935.
28. Canitrot Y, Falinski R, Louat T, et al. p210 BCR/ABL kinase regulates nucleotide excision repair (NER) and resistance to UV radiation. *Blood* 2003; 102: 2632–2637.
29. Sliwinski T, Czechowska A, Szemraj J, et al. STI571 reduces NER activity in BCR/ABL-expressing cells. *Mutat Res* 2008; 654: 162–167.
30. Stoklosa T, Poplawski T, Koptyra M, et al. BCR/ABL inhibits mismatch repair to protect from apoptosis and induce point mutations. *Cancer Res* 2008; 68: 2576–2580.
31. Sawyers CL. Opportunities and challenges in the development of kinase inhibitor therapy for cancer. *Genes Dev* 2004; 17: 2998–3010.
32. Lombardo LJ, Lee FY, Chen P, et al. Discovery of N-(2-chloro-6-methyl-phenyl)-2-(6-(4-(2-hydroxyethyl)-piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. *J Med Chem* 2004; 47: 6658–6661.
33. Shah NP, Tran C, Lee FY, et al. Overriding imatinib resistance with a novel ABL kinase inhibitor. *Science* 2004; 305: 399–401.
34. Le Coutre P, Schwarz M, Kim TD. New developments in tyrosine kinase inhibitor therapy for newly diagnosed chronic myeloid leukemia. *Clin Cancer Res* 2010; 16: 1771–1780.
35. Nam S, Kim D, Cheng JQ, et al. Action of the Src family kinase inhibitor, dasatinib (BMS-354825), on human prostate cancer cells. *Cancer Res* 2005; 65: 9185–9189.
36. Schittenhelm MM, Shiraga S, Schroeder A, et al. Dasatinib (BMS-354825), a dual SRC/ABL kinase inhibitor, inhibits the kinase activity of wild-type, juxtamembrane, and activation loop mutant KIT isoforms associated with human malignancies. *Cancer Res* 2006; 66: 473–481.
37. Tokarski JS, Newitt JA, Chang CY, et al. The structure of dasatinib (BMS-354825) bound to activated ABL kinase domain elucidates its inhibitory activity against imatinib-resistant ABL mutants. *Cancer Res* 2006; 66: 5790–5797.
38. O'Hare T, Walters DK, Stoffregen EP, et al. In vitro activity of Bcr-Abl inhibitors AMN107 and BMS-354825 against clinically relevant imatinib-resistant Abl kinase domain mutants. *Cancer Res* 2005; 66: 4500–4505.
39. Ramirez P, DiPersio JF. Therapy options in imatinib failures. *Oncologist* 2008; 13: 424–434.
40. Soverini S, Colarossi S, Gnani A, et al. Resistance to dasatinib in Philadelphia-positive leukemia patients and the presence or the selection of mutations at residues 315 and 317 in the BCR-ABL kinase domain. *Hematologica* 2007; 92: 401–404.
41. Khorashad JS, Milojkovic D, Mehta P, et al. In vivo kinetics of kinase domain mutations in CML patients treated with dasatinib after failing imatinib. *Blood* 2008; 111: 2378–2381.
42. Burgess MR, Skaggs BJ, Shah NP, et al. Comparative analysis of two clinically active BCR-ABL kinase inhibitors reveals the role of conformation-specific binding in resistance. *Proc Natl Acad Sci U S A* 2005; 102: 3395–3400.
43. Soverini S, Martinelli G, Colarossi S, et al. Presence or the emergence of a F317L BCR-ABL mutation may be associated with resistance to dasatinib in Philadelphia chromosome-positive leukemia. *J Clin Oncol* 2006; 24: e51–e52.
44. Shah NP, Skaggs BJ, Branford S, et al. Sequential ABL kinase inhibitor therapy selects for compound drug-resistant BCR-ABL mutations with altered oncogenic potency. *J Clin Invest* 2007; 117: 2562–2569.
45. Lee F, Fandi A, Voi M. Overcoming kinase resistance in chronic myeloid leukemia. *Int J Biochem Cell Biol* 2008; 40: 334–343.
46. Weisberg E, Manley PW, Breitenstein W, et al. Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl [published correction appears in *Cancer Cell* 2005; 7: 399]. *Cancer Cell* 2005; 7: 129–141.
47. Stover EH, Chen J, Lee BH, et al. The small molecule tyrosine kinase inhibitor AMN107 inhibits TEL-PDGFRbeta and FIP1L1-PDGFRalpha in vitro and in vivo. *Blood* 2005; 106: 3206–3213.
48. Manley PW, Cowan-Jacob SW, Fendrich G, et al. Molecular interactions between the highly selective

pan-Bcr-Abl inhibitor, AMN107, and the tyrosine kinase domain of Abl. *Blood* 2005; 106: abstract 3365.

49. Golemovic M, Verstovsek S, Giles F, et al. AMN107, a novel aminopyrimidine inhibitor of Bcr-Abl, has in vitro activity against imatinib-resistant chronic myeloid leukemia. *Clin Cancer Res* 2005; 11: 4941–4947.
50. Kantarjian H, Giles F, Wunderle L, et al. Nilotinib in imatinib-resistant CML and Philadelphia chromosome-positive ALL. *N Engl J Med* 2006; 354: 2542–2551.
51. Deininger MW. Nilotinib. *Clin Cancer Res* 2008; 14: 4027–4031.
52. O'Hare T, Eide CA, Tyner JW, et al. SGX393 inhibits the CML mutant Bcr-AblT315I and preempts in vitro resistance when combined with nilotinib or dasatinib. *Proc Natl Acad Sci U S A* 2008; 105: 5507–5512.
53. Quintás-Cardama A, Kantarjian H, Cortes J. Flying under the radar: the new wave of BCR-ABL inhibitors. *Nat Rev Drug Discov* 2007; 6: 834–848.
54. Quintás-Cardama A, Cortes J. Therapeutic options against BCR-ABL1 T315I-positive chronic myelogenous leukemia. *Clin Cancer Res* 2008; 14: 4392–4399.
55. Zhang J, Adrián FJ, Jahnke W, et al. Targeting Bcr-Abl by combining allosterix with ATP-binding-site inhibitors. *Nature* 2010; 463: 501–506.
56. Anderson SM, Hayward WS, Neel BG, et al. Avian erythroblastosis virus produces two mRNAs. *J Virol* 1980; 36: 676–683.
57. Vennström B, Bishop JM. Isolation and characterization of chicken DNA homologous to the two putative oncogenes of avian erythroblastosis virus. *Cell* 1982; 28: 135–143.
58. Yamamoto T, Hihara H, Nishida T, et al. A new avian erythroblastosis virus, AEV-H, carries erbB gene responsible for the induction of both erythroblastosis and sarcomas. *Cell* 1983; 34: 225–232.
59. Velu TJ, Beguinot L, Vass WC, et al. Epidermal-growth-factor-dependent transformation by a human EGF receptor proto-oncogene. *Science* 1987; 238: 1408–1410.
60. Hudziak RM, Schlessinger J, Ullrich A. Increased expression of the putative growth factor receptor p185HER2 causes transformation and tumorigenesis of NIH 3T3 cells. *Proc Natl Acad Sci U S A* 1987; 84: 7159–7163.
61. Pao W, Miller VA, Kris MG. 'Targeting' the epidermal growth factor receptor tyrosine kinase with gefitinib (Iressa) in non-small cell lung cancer (NSCLC). *Sem Cancer Biol* 2004; 14: 33–40.
62. Ward WH, Cook PN, Slater AM, et al. Epidermal growth factor receptor tyrosine kinase: investigation of catalytic mechanism, structure-based searching and discovery of a potent inhibitor. *Biochem Pharmacol* 1994; 48: 659–666.
63. Sirotnak FM. Studies with ZD1839 in preclinical models. *Semin Oncol* 2003; 30: 12–20.
64. Wakeling AE, Guy SP, Woodburn JR, et al. ZD1839 (Iressa): an orally active inhibitor of epidermal growth factor signaling with potential for cancer therapy. *Cancer Res* 2002; 62: 5749–5754.
65. Magne N, Fischel JL, Dubreuil A, et al. Influence of epidermal growth factor receptor (EGFR), p53 and intrinsic MAP kinase pathway status of tumour cells on the antiproliferative effect of ZD1839 ("Iressa"). *Br J Cancer* 2002; 86: 1518–1523.
66. Di Gennaro E, Barbarino M, Bruzzese F, et al. Critical role of both p27KIP1 and p21CIP1/WAF1 in the antiproliferative effect of ZD1839 ('Iressa'), an epidermal growth factor receptor tyrosine kinase inhibitor, in head and neck squamous carcinoma cells. *J Cell Physiol* 2003; 195: 139–150.
67. Ciardiello F, Caputo R, Bianco R, et al. Inhibition of growth factor production and angiogenesis in human cancer cells by ZD1839 (Iressa), a selective epidermal growth factor receptor tyrosine kinase inhibitor. *Clin Cancer Res* 2001; 7: 2958–2970.
68. Hirata A, Ogawa S, Kometani T, et al. ZD1839 (Iressa) induces antiangiogenic effects through inhibition of epidermal growth factor receptor tyrosine kinase. *Cancer Res* 2002; 62: 2554–2560.
69. Ono M, Kuwano M. Molecular mechanisms of epidermal growth factor receptor (EGFR) activation and response to gefitinib and other EGFR-targeted drugs. *Clin Cancer Res* 2006; 12: 7242–7251.
70. Fukui T, Mitsudomi T. Mutations in the epidermal growth factor receptor gene and effects of EGFR-tyrosine kinase inhibitors on lung cancers. *Gen Thorac Cardiovasc Surg* 2008; 56: 97–103.
71. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004; 350: 3129–3139.
72. Paez JG, Jänne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004; 304: 1497–1500.
73. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci U S A* 2004; 101: 13306–13311.
74. Irmer D, Funk JO, Blaukat A. EGFR kinase domain mutations: functional impact and relevance for lung cancer therapy. *Oncogene* 2007; 26: 5693–5701.
75. Riely GJ, Pao W, Pham D, et al. Clinical course of patients with non-small cell lung cancer and epidermal growth factor receptor exon 19 and exon 21 mutations treated with gefitinib or erlotinib. *Clin Cancer Res* 2006; 12: 839–844.
76. Jänne PA, Johnson BE. Effect of epidermal growth factor receptor tyrosine kinase domain mutations on the outcome of patients with non-small cell lung cancer treated with epidermal growth factor receptor tyrosine kinase inhibitors. *Clin Cancer Res* 2006; 12(14 part 2): 4416S–4420S.
77. Cappuzzo F, Magrini E, Ceresoli GL, et al. Akt phosphorylation and gefitinib efficacy in patients with advanced non-small-cell lung cancer. *J Natl Cancer Inst* 2004; 96: 1133–1141.
78. Ono M, Hirata A, Kometani T, et al. Sensitivity to gefitinib (Iressa, ZD1839) in non-small cell lung cancer cell lines correlates with dependence on the epidermal growth factor (EGF) receptor/extracellular signal-regulated kinase 1/2 and EGF receptor/Akt pathway for proliferation. *Mol Cancer Ther* 2004; 3: 465–472.
79. Han SW, Hwang PG, Chung DH, et al. Epidermal growth factor receptor (EGFR) downstream molecules as response predictive markers for gefitinib (Iressa,

ZD1839) in chemotherapy-resistant non-small cell lung cancer. *Int J Cancer* 2005; 113: 109–115.

80. Pollack VA, Savage DM, Baker DA, et al. Inhibition of epidermal growth factor receptor-associated tyrosine phosphorylation in human carcinomas with CP-358,774: dynamics of receptor inhibition in situ and antitumor effects in athymic mice. *J Pharmacol Exp Ther* 1999; 291: 739–748.
81. Lal A, Glazer CA, Martinson HM, et al. Mutant epidermal growth factor receptor up-regulates molecular effectors of tumor invasion. *Cancer Res* 2002; 62: 3335–3339.
82. Bareschino MA, Schettino C, Troiani T, et al. 2007. Erlotinib in cancer treatment. *Ann Oncol* 2007; 18(suppl): vi35–vi41.
83. Gridelli C, Bareschino MA, Schettino C, et al. Erlotinib in non-small cell lung cancer treatment: current status and future development. *Oncologist* 2007; 12: 840–849.
84. Kobayashi S, Boggon TJ, Dayaram T, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2005; 352: 786–792.
85. Pao W, Miller VA, Politi KA, et al. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PloS Med* 2005; 3: 225–235.
86. Bell DW, Gore I, Okimoto RA, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. *Nat Genet* 2005; 37: 1315–1316.
87. Godin-Heymann N, Bryant I, Rivera MN, et al. Oncogenic activity of epidermal growth factor receptor kinase mutant alleles is enhanced by the T790M drug resistance mutation. *Cancer Res* 2007; 67: 7319–7326.
88. Yun CH, Mengwasser KE, Toms AV, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci U S A* 2008; 105: 2070–2075.
89. Kosaka T, Yatabe Y, Endoh H, et al. Analysis of epidermal growth factor receptor gene mutation in patients with non-small cell lung cancer and acquired resistance to gefitinib. *Clin Cancer Res* 2006; 12: 5764–5769.
90. Balak MN, Gong Y, Riely GJ, et al. Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors. *Clin Cancer Res* 2006; 12: 6494–6501.
91. Pao W, Wang TY, Riely GJ, et al. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med* 2005; 2: 57–61.
92. Engelman JA, Zejnullahu K, Mitsudomi T, et al. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 2007; 316: 1039–1043.
93. Bean J, Brennan C, Shih JY, et al. MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. *Proc Natl Acad Sci U S A* 2007; 104: 20932–20937.
94. Engelman JA, Mukohara T, Zejnullahu K, et al. Allelic dilution obscures detection of a biologically significant resistance mutation in EGFR-amplified lung cancer. *J Clin Invest* 2006; 116: 2695–2706.
95. Morgillo F, Woo JK, Kim ES, et al. Heterodimerization of insulin-like growth factor receptor epidermal growth factor receptor and induction of survivin expression counteract the antitumor action of erlotinib. *Cancer Res* 2006; 66: 10100–10111.
96. Guix M, Faber AC, Wang SE, et al. Acquired resistance to EGFR tyrosine kinase inhibitors in cancer cells is mediated by loss of IGF-binding proteins. *J Clin Invest* 2008; 118: 2609–2619.
97. Mueller KL, Hunter LA, Ethier SP, et al. Met and c-Src cooperate to compensate for loss of epidermal growth factor receptor kinase activity in breast cancer cells. *Cancer Res* 2008; 68: 3314–3322.
98. Rabindran SK, Discafani CM, Rosfjord EC, et al. Antitumor activity of HKI-272, an orally active, irreversible inhibitor of the HER-2 tyrosine kinase. *Cancer Res* 2004; 64: 3958–3965.
99. Kwak EL, Sordella R, Bell DW, et al. Irreversible inhibitors of the EGF receptor may circumvent acquired resistance to gefitinib. *Proc Natl Acad Sci U S A* 2005; 102: 7665–7670.
100. Li D, Shimamura T, Ji H, et al. Bronchial and peripheral murine lung carcinomas induced by T790M-L858R mutant EGFR respond to HKI-272 and rapamycin combination therapy. *Cancer Cell* 2007; 12: 81–93.
101. Godin-Heymann N, Ulkus L, Brannigan BW, et al. The T790M "gatekeeper" mutation in EGFR mediates resistance to low concentrations of an irreversible EGFR inhibitor. *Mol Cancer Ther* 2008; 7: 874–879.
102. Schechter AL, Stern DF, Vaidyanathan L, et al. The neu oncogene: an erb-B related gene encoding a 185,000-Mr tumor antigen. *Nature* 1984; 312: 513–516.
103. King CR, Kraus MH, Aaronson SA. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. *Science* 1985; 229: 974–976.
104. Di Fiore PP, Pierce JH, Fleming TP, et al. Overexpression of the human EGF receptor confers an EGF-dependent transformed phenotype to NIH 3T3 cells. *Cell* 1987; 51: 1063–1070.
105. Bouchard L, Lamarre L, Tremblay PJ, et al. Stochastic appearance of mammary tumors in transgenic mice carrying the mmtv/c-neu oncogene. *Cell* 1989; 57: 931–936.
106. Slamon DJ, Clark GM, Wong SG, et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 1987; 235: 177–182.
107. Berger MS, Locher GW, Saurer S, et al. Correlation of c-erbB-2 gene amplification and protein expression in human breast carcinoma with nodal status and nuclear grading. *Cancer Res* 1988; 48: 1238–1243.
108. Slamon DJ, Godolphin W, Jones LA, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 1989; 244: 707–712.
109. Kruser TJ, Wheeler DL. Mechanisms of resistance to HER family targeting antibodies. *Exp Cell Res* 2010; 316: 1083–1100.
110. Sliwkowski MX, Lofgren JA, Lewis GD, et al. Nonclinical studies addressing the mechanism of action of trastuzumab (Herceptin). *Semin Oncol* 1999; 26(4 suppl 12): 60–70.
111. Harries M, Smith I. The development and clinical use of trastuzumab (Herceptin). *Endocr Relat Cancer* 2002; 9: 75–85.

112. Carter P, Presta L, Gorman CM, et al. Humanization of an anti-p185HER2 antibody for human cancer therapy. *Proc Natl Acad Sci U S A* 1992; 89: 4285–4289.
113. Nahta R, Esteva FJ. HER-2-targeted therapy: lessons learned and future directions. *Clin Cancer Res* 2003; 9: 5078–5084.
114. Junttila TT, Akita RW, Parsons K, et al. Ligand-independent HER2/HER3/PI3K complex is disrupted by trastuzumab and is effectively inhibited by the PI3K inhibitor GDC-0941. *Cancer Cell* 2009; 15: 429–440.
115. Carter CA, Kelly RJ, Giaccone G. Small-molecule inhibitors of the human epidermal receptor family. *Expert Opin Invest Drugs* 2009; 18: 1829–1842.
116. Yakes FM, Chinratanalab W, Ritter CA, et al. Herceptin-induced inhibition of phosphatidylinositol-3 kinase and Akt Is required for antibody-mediated effects on p27, cyclin D1, and antitumor action. *Cancer Res* 2002; 62: 4132–4141.
117. Ye D, Mendelsohn J, Fan Z. Augmentation of a humanized anti-HER2 mAb 4D5 induced growth inhibition by a human-mouse chimeric anti-EGF receptor mAb C225. *Oncogene* 1999; 18: 731–738.
118. Lane HA, Beuvink I, Motoyama AB, et al. ErbB2 potentiates breast tumor proliferation through modulation of p27(Kip1)-Cdk2 complex formation: receptor overexpression does not determine growth dependency. *Mol Cell Biol* 2000; 20: 3210–3223.
119. Le XF, McWatters A, Wiener J, et al. Anti-HER2 antibody and heregulin suppress growth of HER2-overexpressing human breast cancer cells through different mechanisms. *Clin Cancer Res* 2000; 6: 260–270.
120. Neve RM, Sutterlüty H, Pullen N, et al. Effects of oncogenic ErbB2 on G1 cell cycle regulators in breast tumour cells. *Oncogene* 2000; 19: 1647–1656.
121. Marches R, Uhr JW. Enhancement of the $p27^{KIP1}$-mediated antiproliferative effect of trastuzumab (Herceptin) on HER2-overexpressing tumor cells. *Int J Cancer* 2004; 112: 492–501.
122. Petit AM, Rak J, Hung MC, et al. Neutralizing antibodies against epidermal growth factor and ErbB-2/neu receptor tyrosine kinases down-regulate vascular endothelial growth factor production by tumor cells in vitro and in vivo: angiogenic implications for signal transduction therapy of solid tumors. *Am J Pathol* 1997; 151: 1523–1530.
123. Izumi Y, Xu L, di Tomaso E, et al. Tumour biology: herceptin acts as an anti-angiogenic cocktail. *Nature* 2002; 416: 279–280.
124. Molina MA, Codony-Servat J, Albanell J, et al. Trastuzumab (herceptin), a humanized anti-Her2 receptor monoclonal antibody, inhibits basal and activated Her2 ectodomain cleavage in breast cancer cells. *Cancer Res* 2001; 61: 4744–4749.
125. Cobleigh MA, Vogel CL, Tripathy D, et al. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J Clin Oncol* 1999; 17: 2639–2648.
126. Vogel CL, Cobleigh MA, Tripathy D, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J Clin Oncol* 2002; 20: 719–726.
127. Roy V, Perez EA. Beyond trastuzumab: small molecule tyrosine kinase inhibitors in HER-2-positive breast cancer. *Oncologist* 2009; 14: 1061–1069.
128. Lu Y, Zi X, Zhao Y, et al. Insulin-like growth factor-I receptor signaling and resistance to trastuzumab (Herceptin). *J Natl Cancer Inst* 2001; 93: 1852–1857.
129. Nahta R, Takahashi T, Ueno NT, et al. $P27^{kip1}$ down-regulation is associated with trastuzumab resistance in breast cancer cells. *Cancer Res* 2004; 64: 3981–3986.
130. Lu Y, Zi X, Pollak M. Molecular mechanisms underlying IGF-I-induced attenuation of the growth-inhibitory activity of trastuzumab (Herceptin) on SKBR3 breast cancer cells. *Int J Cancer* 2004; 108: 334–341.
131. Nahta R, Yuan LX, Zhang B, et al. Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells. *Cancer Res* 2005; 65: 11118–11128.
132. Leslie NR, Downes CP. PTEN function: how normal cells control it and tumour cells lose it. *Biochem J* 2004; 382(part 1): 1–11.
133. Nagata Y, Lan KH, Zhou X, et al. PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* 2004; 6: 117–127.
134. Bleeker WK, Lammerts van Bueren JJ, van Ojik HH. Dual mode of action of a human anti-epidermal growth factor receptor monoclonal antibody for cancer therapy. *J Immunol* 2004; 173: 4699–4707.
135. Lurje G, Lenz HJ. EGFR signaling and drug discovery. *Oncology* 2009; 77: 400–410.
136. Perez-Torres M, Guix M, Gonzalez A, et al. Epidermal growth factor receptor (EGFR) antibody down-regulates mutant receptors and inhibits tumors expressing EGFR mutations. *J Biol Chem* 2006; 281: 40183–40192.
137. Doody JF, Wang Y, Patel SN, et al. Inhibitory activity of cetuximab on epidermal growth factor receptor mutations in non small cell lung cancers. *Mol Cancer Ther* 2007; 6: 2642–2651.
138. Steiner P, Joynes C, Bassi R. Tumor growth inhibition with cetuximab and chemotherapy in non-small cell lung cancer xenografts expressing wild-type and mutated epidermal growth factor receptor. *Clin Cancer Res* 2007; 13: 1540–1551.
139. Wu M, Rivkin A, Pham T. Panitumumab: human monoclonal antibody against epidermal growth factor receptors for the treatment of metastatic colorectal cancer. *Clin Ther* 2008; 30: 14–30.
140. Amado RG, Wolf M, Peeters M. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. *J Clin Oncol* 2008; 26: 1626–1634.

Bladder Cancer: Current Management and Opportunities for a Personalized Approach

Matthew D. Galsky, MD[1] and Simon J. Hall, MD[2]

[1]Division of Hematology and Medical Oncology, Mount Sinai School of Medicine, New York, NY
[2]Department of Urology, Mount Sinai School of Medicine, New York, NY

OUTLINE

Molecular Pathogenesis of Bladder Cancer
Management and Opportunities for Personalized Medicine
  Non–Muscle-Invasive Bladder Cancer
  Muscle-Invasive Bladder Cancer
  Metastatic Bladder Cancer
Recent Approaches Exploring Personalized Medicine in Bladder Cancer
Neoadjuvant Setting as a Paradigm for Biomarker and Drug Development
Conclusion

ABSTRACT

Urothelial carcinoma of the bladder exists as a spectrum of clinical states ranging from non–muscle-invasive to muscle-invasive to advanced/metastatic disease. Each clinical state is associated with a unique tumor biology, prognosis, and approach to treatment. The field of personalized medicine offers the promise to individualize therapeutic decisions in each clinical state with the goal of preventing or delaying progression to more advanced disease states while optimizing quality of life. This review will discuss bladder cancer as a paradigm for the development of personalized cancer medicine, focusing on recent efforts, current challenges, and future opportunities. ***Mt Sinai J Med 77:587–596, 2010.*** *© 2010 Mount Sinai School of Medicine*

Address Correspondence to:

Matthew D. Galsky
Director, Genitourinary
Medical Oncology
Mount Sinai School of Medicine
New York, NY
Email: matthew.galsky@mssm.edu

Key Words: bladder cancer, muscle-invasive, non–muscle-invasive, personalized medicine, urothelial carcinoma.

Urothelial carcinoma of the bladder is the fifth most common cancer in the United States, with >70,000 new cases and >14,000 deaths each year.[1] Bladder cancer exists as a spectrum of clinical states ranging from non–muscle-invasive to muscle-invasive to advanced/metastatic disease (Figure 1). Each clinical state is associated with a unique tumor biology, prognosis, and approach to treatment. The goal of treatment in each clinical state is to prevent or delay progression to more advanced disease states, and ultimately death, while maintaining optimal quality of life. Because only a subset of patients in each clinical state will achieve these benefits of therapy, yet all patients treated will be exposed to potential adverse events, there are prime opportunities to individualize care throughout the spectrum of disease. This review will discuss the current standard management of bladder cancer and current/future efforts to integrate a personalized approach to treatment.

MOLECULAR PATHOGENESIS OF BLADDER CANCER

Traditionally, localized bladder cancers were termed either superficial or invasive, differentiating cancers that invaded the bladder muscle from those that were confined to the bladder mucosa or submucosa. The term noninvasive is now used to delineate the heterogeneous behavior and sometimes aggressive features of subsets of these tumors, not represented by the term "superficial." Multiple interacting molecular pathways have been demonstrated to play a role in the genesis of bladder cancer. Noninvasive tumors involve 2 distinct pathways that appear to mediate the different disease phenotypes: those that tend to recur

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20224

© 2010 Mount Sinai School of Medicine

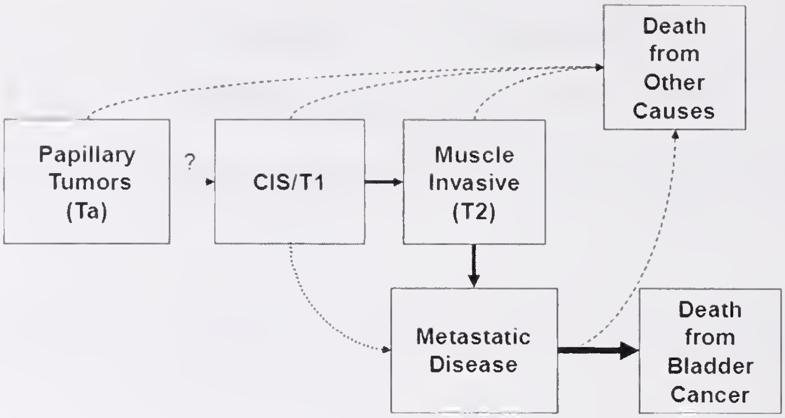


Fig 1. Clinical states representing the natural history of bladder cancer. **Abbreviations:** CIS, carcinoma in situ; T1, tumor extending into the lamina propria but not the muscularis propria; Ta, papillary tumor located above the basement membrane.

but not invade the bladder wall, and those that recur but invade the bladder wall and metastasize. Tumors that are confined to the mucosa of the bladder lining and are low-grade show frequent constitutive activation of the RAS-mitogen-activated protein kinase (RAS-MAPK) pathway, with activating mutations in *HRAS* and *FGFR3* genes. In contrast, high-grade cancers–both those that are noninvasive, carcinoma in situ (CIS) and pT1 that invade the lamina propria, and those that invade the muscle– demonstrate alterations in *TP53* and *RB* genes and pathways.[2] Additional mutations, known to play a role in driving growth and progression of other solid tumors, have recently been identified in small subsets of invasive bladder tumors, including *BRAF*, phosphoinositide 3-kinase, and *AKT*.[3] In addition, loss of the tumor suppressor phosphatase and tensin homolog has been implicated in the pathogenesis of bladder cancer in mouse models.[4] Identification of these altered pathways permits the opportunity to begin to pair patients with these particular mutations with appropriate inhibitors of the relevant aberrant pathways.

MANAGEMENT AND OPPORTUNITIES FOR PERSONALIZED MEDICINE

Non–Muscle-Invasive Bladder Cancer

Non–muscle-invasive bladder cancer is a significant public health problem. Of the approximately 60,000 new cases of bladder cancer per year in the United States, 70%–80% are non–muscle-invasive bladder cancers. Bladder cancer is the fourth most common cancer in men and the eighth most common cancer in women worldwide. The economic burden of this disease is significant, and given the need for frequent cystoscopies, transurethral resections, and intravesical therapy, the estimated lifetime cost of care is $65,000.[5]

Non–muscle-invasive bladder tumors are staged as Tis (CIS), Ta (papillary tumor located above the basement membrane), or T1 (tumor extending into the lamina propria but not the muscularis propria). The clinical behavior and molecular pathogenesis of these lesions are disparate. Both subtypes tend to recur in up to 50% of patients, but low-grade papillary (Ta) tumors rarely progress to invasion, whereas CIS[6] and T1 tumors,[7] both usually high grade, are at significant risk for progression to muscle-invasive disease and metastasis.

The treatment of non–muscle-invasive bladder tumors is based on risk stratification using known prognostic factors that often divide patients into 2 groups: (1) low risk, including low-grade papillary tumors, tumors that are small in size and number, and tumors with a long interval between recurrences; and (2) high risk, including high-grade diffuse CIS, T1 tumors, large tumors, multiple tumors, and short intervals between recurrences. Low-risk tumors are generally managed by a complete transurethral resection of bladder tumor (TURBT) through the cystoscope followed by surveillance. Patients who do recur are treated with intravesicle chemotherapy such as mitomycin. High-risk tumors are generally managed by TURBT plus intravesical instillation of Bacillus Calmette-Guérin (BCG) as opposed to chemotherapy. Given the significant risk of high-grade T1 tumors for progression, especially those that recur following a course of BCG therapy, some urologic oncologists advocate early radical cystectomy for these patients.[8]

Given the public health burden of non–muscle-invasive bladder cancer, there are several potential

DOI:10.1002/MSJ

opportunities for a personalized approach. For example, a common surveillance strategy after treatment often involves cystoscopy and urine cytology every 3 months × 2 years, followed by every 6 months × 2 years, followed by annually. Reliable blood or urine markers to detect recurrent disease in patients undergoing surveillance after initial treatment for non–muscle-invasive disease could spare the cost and hassle of frequent cystoscopic examinations.

> *Given the public health burden of non–muscle-invasive bladder cancer, there are several potential opportunities for a personalized approach.*

Unfortunately, in a recent prospective study of 200 patients with non–muscle-invasive bladder cancer undergoing surveillance with cystoscopy, urine cytology, NMP22 BladderChek (a test for nuclear matrix protein 22), and FISH Urovysion (a urine test designed to detect aneuploidy for chromosomes 3, 7, and 17, and loss of the 9p21 locus via fluorescence in situ hybridization [FISH]) testing, cystoscopy alone remained the most cost-effective strategy to detect recurrence and resulted in the lowest number of cancers missed.[9] The addition of urinary markers added to cost, without increase in tumor detection. Additional critical opportunities for the development of markers for individualized treatment include identification of patients with non–muscle-invasive disease at highest risk for progression and suitable for early cystectomy, identification of patients for treatment with novel targeted agents (eg, fibroblast growth factor receptor inhibitors in patients with *FGFR3* mutations), and selection of patients most likely to respond to BCG.

Muscle-Invasive Bladder Cancer

Radical cystectomy with pelvic lymph node dissection is the standard approach to the management of patients with bladder tumors that infiltrate the muscularis propria (muscle-invasive disease). A radical cystectomy involves a wide excision of the bladder with all of the perivesical fat and tissue in an attempt to achieve negative surgical margins. In men, a prostatectomy is also performed; in women, the urethra, uterus, fallopian tubes, ovaries, anterior vaginal wall, and surrounding fascia are removed. A standard pelvic lymph node dissection includes removal of the distal common iliac, external iliac, obturator, and hypogastric nodes.

The likelihood of cure after radical cystectomy for muscle-invasive disease has been associated with the pathologic stage, the "quality" of surgery (including margin status and extent of lymphadenectomy), and the use of perioperative chemotherapy. In a large single-center series including 1054 patients who underwent radical cystectomy, the 5-year disease-free survival for patients with node-negative organ-confined disease exceeded 80% but dropped to 62%–78% for patients with extravesical disease (pT3) and 50% for patients with invasion of local organs (pT4). Patients with lymph node involvement had worse outcomes, with a 5-year disease-free survival of 35%.[10]

Given the poor outcomes in patients with extravesical extension and/or regional lymph node involvement with surgery alone, several studies have explored the use of perioperative chemotherapy in an attempt to eradicate micrometastases and improve the likelihood of cure. Neoadjuvant therapy refers to chemotherapy administered prior to definitive local therapy. Though many neoadjuvant studies have suffered from small sample size and use of suboptimal chemotherapy regimens, 2 large studies have established neoadjuvant cisplatin-based chemotherapy as a standard approach in muscle-invasive bladder cancer. A study led by the Southwest Oncology Group randomized patients with muscle-invasive bladder cancer to radical cystectomy alone (154 patients) compared with 3 cycles of methotrexate, vinblastine, doxorubicin, and cisplatin (MVAC) followed by radical cystectomy (153 patients).[11] The integration of neoadjuvant chemotherapy was associated with a higher rate of complete pathologic response (pT0 38% compared with 15%, $P < 0.001$). At a median follow-up of 8.7 years, improvements in median survival (77 compared with 46 months, $P = 0.06$) and 5-year survival (57% compared with 43%, $P = 0.06$) favored the use of neoadjuvant MVAC. Notably, the enhancement in survival was correlated with achieving pT0 status after chemotherapy. A European trial randomized 976 patients with muscle-invasive bladder cancer to neoadjuvant cisplatin, methotrexate, and vinblastine (CMV; 491 patients) versus local therapy alone (485 patients).[12] Management of the primary tumor involved cystectomy, radiation therapy, or both. At a median follow-up of approximately 7 years, an improvement in survival was observed for patients who received neoadjuvant chemotherapy ($P = 0.048$; hazard ratio [HR]: 0.85; 95% confidence interval [CI]: 0.72–1.0).[13] Trials exploring adjuvant chemotherapy (chemotherapy given after definitive local therapy) have yielded inconclusive results, but have been small and underpowered.

As noted, approximately 50% of patients with muscle-invasive bladder cancer are cured with surgery alone, and the absolute survival benefit with the integration of perioperative chemotherapy is approximately 5%–10%. Therefore, for every 100 patients treated with chemotherapy + surgery for muscle-invasive disease, 50 patients would have been cured with surgery alone and are given chemotherapy unnecessarily, with another 40 to 45 patients not cured despite undergoing both treatments. This uncertainty regarding which patients are benefiting from the integration of systemic therapy likely contributes to the poor uptake of neoadjuvant/adjuvant chemotherapy in the urologic community.[14] There is a crucial need for personalized approaches to (1) identify patients who require integration of systemic therapy, and (2) to select patients likely to benefit from integration of a particular systemic regimen. Clinical and pathologic prognostic factors have enhanced the ability to predict patients likely to be cured with surgery alone beyond traditional pathologic staging. For example, a multinational database was utilized to construct a nomogram (available at www.nomograms.org) to provide individualized risk predictions.[15] The addition of other readily measurable pathologic features, such as lymphovascular invasion, may further refine such tools.[16] Several studies have explored genes and gene panels in an attempt to improve prognostication.[17,18] Overexpression of 4 genes (JUN, MAP2K6, STAT3, and ICAM1), molecules associated with cell survival (Bcl-2, caspase-3, p53, survivin), and insensitivity to antigrowth signals (p53, p21, p16, pRB) are correlated with poor outcomes following cystectomy.[19–21] However, unlike the gene panels developed for prognostication in clinically localized breast cancer (eg, Oncotype DX, MammaPrint), these tools have not been extensively validated or routinely integrated into clinical practice.

> *There is a crucial need for personalized approaches to (1) identify patients who require integration of systemic therapy, and (2) to select patients likely to benefit from integration of a particular systemic regimen.*

In an attempt to improve functional outcomes and avoid the significant morbidity of a cystectomy, yet maintain the likelihood of cure, several bladder-sparing approaches have been explored for bladder cancer, including partial cystectomy and a multimodality approach including TURBT, chemotherapy, and radiation. There have been no randomized trials comparing the results of bladder-sparing therapy with radical cystectomy. However, compared with historical controls, the use of a combined modality approach in appropriately selected patients (ie, a solitary tumor that can be completely resected with TURBT, lack of multifocal CIS, lack of involvement of the ureteral orifice) has yielded similar outcomes. Large phase II studies exploring combined modality therapy have reported 5-year survival rates of 50%–60%, with approximately 70% of those patients maintaining an intact bladder.[22] Given that up to 25% of patients treated with a bladder-sparing approach will ultimately require cystectomy, and approximately 50% of patients will ultimately develop recurrence, key opportunities exist for the development of prognostic and predictive markers to select ideal patients for bladder preservation beyond relying on clinical features alone.

Metastatic Bladder Cancer

Approximately 25% of patients with bladder cancer present de novo with metastatic disease, and approximately 50% of patients undergoing potentially curative surgery will develop metastatic recurrence. Bladder cancer is a chemosensitive neoplasm characterized by relatively high response rates to multiple chemotherapeutic agents compared with other solid tumors. However, with the exception of a small subset of patients with lymph node–only metastases or oligometastatic disease who may achieve durable disease control with a combination of chemotherapy and surgery, the majority of patients with bladder cancer are not curable with conventional treatments.

The MVAC regimen, developed in the 1980s, represented a step forward in the management of patients with metastatic bladder cancer. In the initial phase II studies with this regimen, objective responses (a decrease in tumor size by at least 30% on imaging) were observed in 71% of patients, with complete responses observed in 50% of patients.[23] Several larger trials subsequently confirmed the activity of MVAC, albeit with lower response rates. Subsequently, randomized trials proved MVAC as superior in efficacy to single-agent cisplatin, and the combination regimen of cisplatin, cyclophosphamide, and cisplatin with MVAC was adopted as a standard of care (Table 1). However, MVAC was associated with significant limitations: the median survival with this regimen was approximately 12.5 months, and treatment-related toxicities included febrile

Table 1. *Randomized Trials in Metastatic Bladder Cancer.*

Regimen	No. of Patients	OR (%)	MS (mo)	*P* Value
MVAC[46]	246	36	12.5	<0.0002
Cisplatin		11	8.2	
MVAC[47]	110	65	12.6	<0.05
CISCA		46	10	
MVAC[48]	169	59	12.5	0.17
FAP		42	12.5	
MVAC[49]	263	58	14.1	0.122
HD-MVAC		72	15.5	
MVAC[24]	408	46	14.8	0.746
Gemcitabine + cisplatin		50	13.8	
MVAC[50]	220	54	14.2	0.025
Docetaxel + cisplatin		37	9.3	
MVAC[51]*	85	40	14.2	0.41
Paclitaxel + carboplatin		28	13.8	
Gemcitabine + cisplatin[25]	627	46	12.8	NS
Gemcitabine + cisplatin + paclitaxel		57	15.7	

Abbreviations: CISCA, cyclophosphamide, cisplatin, doxorubicin; CMV, cisplatin, methotrexate, vinblastine; CR, complete response; FAP, 5-fluorouracil, interferon-alpha-2b, cisplatin; HD-MVAC, high-dose MVAC; MS, median survival; MV, methotrexate, vinblastine; MVAC, methotrexate, vinblastine, doxorubicin, cisplatin; NS, not significant; OR, overall response.

*Trial closed early with only 85 patients.

Table 2. *Second-Line Chemotherapy in Metastatic Bladder Cancer.*

Reference	Regimen	No. of Patients	OR	95% CI
Vaughn *et al.*[52]	Paclitaxel	31	10%	0%–26%
McCaffrey *et al.*[53]	Docetaxel	30	13%	4%–30%
Witte *et al.*[54]	Ifosfamide	56	20%	10%–32%
Lorusso *et al.*[55]	Gemcitabine	35	20%	10%–36%
Vaishampayan *et al.*[56]	Paclitaxel + carboplatin	44	16%	7%–30%
Culine *et al.*[57]	Vinflunine	51	18%	8%–31%
Galsky *et al.*[58]	Pemetrexed	13	8%	NR
Sweeney *et al.*[59]	Pemetrexed	47	28%	16%–43%

Abbreviations: CI, confidence interval; OR, overall response; NR, not reported.

neutropenia, mucositis, renal impairment, hearing loss, and peripheral neuropathy.

In an effort to improve outcomes and tolerability, newer chemotherapeutic agents developed over the past 2 decades have been explored in patients with metastatic bladder cancer. A landmark phase III trial randomized 405 patients with metastatic urothelial carcinoma to gemcitabine + cisplatin (GC) versus MVAC.[24] Although this trial was not designed as a noninferiority trial, the median overall survival was similar in both arms (MVAC 15.2 months, GC 14 months; HR: 1.04, 95% CI: 0.82–1.32; $P = 0.75$). In addition, the GC arm was associated with a favorable toxicity profile compared with MVAC: neutropenic fever (2% versus 14%), neutropenic sepsis (1% versus 12%), and grade 3 or 4 mucositis (1% versus 22%) and alopecia (11% versus 55%). Based on this trial, GC has become a widely utilized treatment standard for metastatic urothelial carcinoma. Notably, the GC regimen represents an advance in terms of tolerability; however, despite attempts to build on this regimen,[25] there have been no improvements in efficacy since the MVAC regimen was developed 30 years ago, and novel approaches are clearly needed. Though several agents have mild to moderate activity in patients with bladder cancer that has progressed on first-line chemotherapy, there are no agents approved by the US Food and Drug Administration for second-line therapy (Table 2). A recent randomized phase III trial failed to demonstrate improved survival with the novel vinca alkaloid vinflunine compared with placebo in this patient population.[26]

Several pathways central to cancer growth and progression are important in the pathogenesis of bladder cancer and represent attractive targets for therapeutic intervention. Thus far, inhibition of tumor angiogenesis and inhibition of signal transduction through the epidermal growth factor family of receptors (EGFR) have been most extensively studied in the clinic. Preclinical studies have demonstrated the antitumor activity of the EGFR tyrosine kinase inhibitor gefitinib in bladder cancer cell lines and correlated activity with the degree of expression of EGFR.[27] However, in a clinical trial of patients with bladder cancer that had progressed on first-line chemotherapy, gefitinib was associated with an objective response rate of only 3%.[28] Although pretreatment tumor tissue for EGFR expression was mandatory for enrollment on this trial, enrollment was not restricted to patients with EGFR overexpression.

> *Several pathways central to cancer growth and progression are important in the pathogenesis of bladder cancer and represent attractive targets for therapeutic intervention.*

Several trials exploring targeting EGFR in combination with conventional chemotherapy are either completed or ongoing, and results available thus far have been disappointing. Notably, preclinical studies suggest cells that have undergone epithelial-to-mesenchymal transition may be resistant to EGFR inhibition, which may in part explain the clinical results with this class of agents.[29]

Targeting the tumor neovasculature has offered more promising preliminary results. Sunitinib, an oral inhibitor of several tyrosine kinases including the vascular endothelial growth factor receptor and platelet derived growth factor receptor, was explored in a phase II trial of heavily pretreated patients with metastatic bladder cancer. Sunitinib was associated with a "clinical benefit rate" (partial response + stable disease [SD] at 3 months) in 31% of 41 patients.[30] Based on preclinical data supporting combination therapy with sunitinib plus chemotherapy,[31] the combination of gemcitabine, cisplatin, and sunitinib has been explored in chemonaïve patients with metastatic bladder cancer.[32] Preliminary results with this regimen indicate promising activity, but also additive hematologic toxicities resulting in difficulty administrating adequate doses of all agents. The combination of GC with the antivascular endothelial growth factor monoclonal antibody bevacizumab has shown promising response outcomes in a large phase II trial.[33] This regimen has been moved forward to definitive testing in a randomized phase III trial led by the Cancer and Leukemia Group B.

Given that multiple chemotherapeutic agents, and "targeted" therapies, exhibit antitumor activity only in a subset of patients with bladder cancer, yet all patients exposed to these drugs are at risk for treatment-related side effects, markers predictive of response to therapy represent a major goal for personalized cancer medicine.

Given that multiple chemotherapeutic agents, and "targeted" therapies, exhibit antitumor activity only in a subset of patients with bladder cancer, yet all patients exposed to these drugs are at risk for treatment-related side effects, markers predictive of response to therapy represent a major goal for personalized cancer medicine.

RECENT APPROACHES EXPLORING PERSONALIZED MEDICINE IN BLADDER CANCER

As outlined above, there are multiple opportunities for personalized medicine to optimize care of patients throughout the spectrum of clinical states of bladder cancer. Clinical trials have just begun to explore this approach in patients. For patients with clinically localized bladder cancer, prognostic biomarkers to select patients destined for systemic relapse are highly sought after. A retrospective analysis demonstrated that immunohistochemical (IHC) detection of p53 (a surrogate for mutant p53 status given the longer half-life than the wild-type protein) significantly correlated with a higher likelihood of recurrence after cystectomy.[34] Further data from a retrospective trial of adjuvant cisplatin-based chemotherapy suggested that the benefit of adjuvant chemotherapy was restricted to patients with altered p53 status, presumably due to the use of DNA-damaging agents.[35] These data led to the design of a prospective trial that registered patients with pT1-T2 N0 bladder cancer status post cystectomy who consented for p53 analysis of their tumor specimen.[36] Patients with "altered" p53 (>10% nuclear positivity by IHC) were randomized to 3 cycles of adjuvant MVAC versus observation. Patients with wild-type p53 were observed. This study was designed to detect a 20% absolute decrease in recurrence rate at 3 years. However, the study was closed when the results of a preplanned interim analysis after randomization of 110 patients suggested that the probability of achieving a significant difference in the time to recurrence in the randomized population was highly unlikely. Furthermore, the prognostic significance of p53 testing was not confirmed in this trial, stressing the importance of prospective validation of retrospective studies.

Although there have not yet been prospective studies selecting patients for treatment with specific cytotoxic chemotherapeutic agents based on molecular markers, several retrospective studies support the promise of this approach. Levels of the DNA-repair genes ERCC1, RRM1, BRCA1, and caveolin-1 were evaluated in tumor tissue from 57 patients with bladder cancer treated with cisplatin-based combination chemotherapy.[37] The median survival was significantly higher in patients with low ERCC1 levels (25.4 versus 15.4 months; $P = 0.03$). "Signatures" of response to platinum-based chemotherapy based on gene-expression profiles have been generated and correlated with clinical outcomes in patients with bladder cancer.[17,38,39] A novel approach, known as coexpression extrapolation or COXEN, utilizes publicly available data regarding gene-expression profiling and drug sensitivity from the NCI-60 panel of tumor cell lines as a Rosetta Stone to predict chemosensitivity of gene–expression-profiled bladder cancer samples using a computational algorithm. The COXEN approach has successfully predicted antitumor outcomes in patients with bladder cancer treated with chemotherapy in retrospective studies.[40] A study is currently being planned that utilizes the COXEN algorithm to select from 1 of 3 chemotherapeutic agents for second-line treatment of patients with metastatic bladder cancer.

Several clinical trials have explored inhibitors of the EGFR pathway in selected patients with metastatic bladder cancer. A phase II trial of chemonaïve patients with metastatic bladder cancer explored the combination of cytotoxic chemotherapy (GC + paclitaxel) + the anti-HER-2 monoclonal antibody trastuzumab. Patients were selected for enrollment based HER-2 overexpression by IHC, gene amplification, and/or elevated serum HER-2. Thirty-one of 44 patients (70%) achieved objective responses (5 complete and 26 partial). However, these results are similar to the results achieved with cytotoxic chemotherapy alone, and the contribution of trastuzumab in this single-arm phase II trial is unclear.

A phase II study explored the dual HER-2/EGFR pathway inhibitor lapatinib as second-line therapy in patients with metastatic bladder cancer. Patients were eligible provided that they had 1+, 2+, or 3+ expression of either EFGR or HER-2 by IHC (from either primary or metastatic tumor samples).[41] Fifty-nine patients were enrolled in the study, 25 of whom (42%) could not be evaluated for response. An objective response to treatment was observed in 1.7% (95% CI: 0.0%–9.1%) of patients; however, 18 patients (31%; 95% CI: 19%–44%) achieved SD. Clinical benefit (overall response rate and SD) was found to be correlated with EGFR overexpression, and, to some extent, HER-2 overexpression. Another phase II trial exploring lapatinib utilized a novel trial design focused on the putative predictive biomarker, rather than the specific tumor type, coined the "target specific, histology independent, randomized discontinuation design."[42] Patients enrolled on this study could have a wide variety of metastatic solid tumors (including bladder cancer, endometrial cancer, ovarian cancer, and gastric cancer) provided tumor tissue available for testing demonstrated *HER2* gene amplification by FISH. Unfortunately, this trial was closed early due to poor accrual. However, of the 33 patients with metastatic bladder cancer tested, 12 patients had *HER2* amplification by FISH.[43] There were no objective responses to treatment among this subgroup of bladder cancer patients. These trials, which all utilized inhibitors of the same pathway but different tests and "cut-offs" for the putative predictive biomarkers, highlight the challenges of drug development in the era of targeted therapeutics.

NEOADJUVANT SETTING AS A PARADIGM FOR BIOMARKER AND DRUG DEVELOPMENT

The traditional paradigm of oncologic drug development often involves first evaluating the efficacy of novel therapeutic agents in patients with advanced refractory disease. This setting, however, may not provide an optimal signal of drug activity and generally does not permit access to pre- and post-treatment tissue for biomarker development. Alternatively, evaluation of novel therapeutics in "window of opportunity" trials prior to patients undergoing radical cystectomy of muscle-invasive bladder cancer may facilitate biomarker development, as pre- and post-treatment tissue is readily available. Furthermore, the use of pathologic complete response rate as an intermediate endpoint may accelerate the assessment of antitumor activity. This approach has been used to explore the antitumor and biologic activity of the EGFR inhibitor erlotinib.[44] Among 20 patients with muscle-invasive bladder cancer treated with erlotinib 150 mg

Evaluation of novel therapeutics in "window of opportunity" trials prior to patients undergoing radical cystectomy of muscle-invasive bladder cancer may facilitate biomarker development, as pre- and post-treatment tissue is readily available.

daily × 4 weeks followed by radical cystectomy, 25% achieved a pathologic complete response and 35% of patients were clinically downstaged. However, given the small size of this trial, it is unclear if these results are superior to those achieved with TURBT alone, and the molecular correlates from this study have not yet been reported. The Src inhibitor dasatinib has shown activity in Src-overexpressing human bladder cancer cell lines, and we are currently exploring the antitumor and biologic activity of dasatinib in a "window of opportunity" neoadjuvant trial in patients with bladder cancer.[45] In addition, we are exploring the antitumor and biologic activity of the combination of GC + sunitinib as neoadjuvant therapy in patients with muscle-invasive bladder cancer. In the age of targeted therapeutics, with the need for pre- and post-treatment tissue to determine mechanisms of sensitivity and resistance, this design will likely become more common.

CONCLUSION

Given the heterogeneous clinical stages of bladder cancer, characterized by competing risks of cancer-related death and treatment-related morbidity/mortality, individualized approaches are urgently needed. Attempts to individualize treatment thus far have met with variable success. As the signaling networks that regulate bladder cancer growth and metastases are better understood, this knowledge may translate to approaches that truly match the "right treatment for the right patient at the right time." The goals of intervention in each clinical state (Figure 1) are to prevent or delay transition to the next clinical state while minimizing treatment-related adverse events.

DISCLOSURES

Potential conflict of interest: Dr. Galsky has received speaking fees from Novartis.

REFERENCES

1. Jemal A, Siegel R, Xu J, et al. Cancer statistics, 2010. *CA Cancer J Clin* 2010; 60: 277–300.
2. Mitra AP, Datar RH, Cote RJ. Molecular pathways in invasive bladder cancer: new insights into mechanisms, progression, and target identification. *J Clin Oncol* 2006; 24: 5552–5564.
3. Iyer G, Lin O, Al-Ahmadie H, et al. Identification and characterization of FGFR3 mutations in high-grade bladder cancer. ASCO Genitourinary Cancers Symposium 2010: Abstract 296.
4. Puzio-Kuter AM, Castillo-Martin M, Kinkade CW, et al. Inactivation of p53 and Pten promotes invasive bladder cancer. *Genes Dev* 2009; 23: 675–680.
5. Avritscher EB, Cooksley CD, Grossman HB, et al. Clinical model of lifetime cost of treating bladder cancer and associated complications. *Urology* 2006; 68: 549–553.
6. Lamm DL. Carcinoma in situ. *Urol Clin North Am* 1992; 19: 499–508.
7. Anderström C, Johansson S, Nilsson S. The significance of lamina propria invasion on the prognosis of patients with bladder tumors. *J Urol* 1980; 124: 23–26.
8. Shahin O, Thalmann GN, Rentsch C, et al. A retrospective analysis of 153 patients treated with or without intravesical bacillus Calmette-Guerin for primary stage T1 grade 3 bladder cancer: recurrence, progression and survival. *J Urol* 2003; 169: 96–100.
9. Karam J, Shah J, Kader AK, et al. Prospective trial to identify optimal bladder cancer surveillance protocol: reducing costs while maximizing sensitivity. ASCO Genitourinary Cancers Symposium 2010: Abstract 275.
10. Stein JP, Lieskovsky G, Cote R, et al. Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. *J Clin Oncol* 2001; 19: 666–675.
11. Grossman HB, Natale RB, Tangen CM, et al. Neoadjuvant chemotherapy plus cystectomy compared with cystectomy alone for locally advanced bladder cancer. *N Engl J Med* 2003; 349: 859–866.
12. Neoadjuvant cisplatin, methotrexate, and vinblastine chemotherapy for muscle-invasive bladder cancer: a randomised controlled trial. International Collaboration of Trialists [published correction appears in *Lancet* 1999;354:1650]. *Lancet* 1999; 354: 533–540.
13. Hall RR. Updated results of a randomized controlled trial of neoadjuvant cisplatin, methotrexate and vinblastine chemotherapy for muscle-invasive bladder cancer. *Proc Am Soc Clin Oncol* 2002; 21: Abstract 710.
14. David KA, Milowsky MI, Ritchey J, et al. Low incidence of perioperative chemotherapy for stage III bladder cancer 1998 to 2003: a report from the National Cancer Data Base. *J Urol* 2007; 178: 451–454.
15. Bochner BH, Kattan MW, Vora KC. Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer. *J Clin Oncol* 2006; 24: 3967–3972.
16. Shariat SF, Svatek RS, Tilki D, et al. International validation of the prognostic value of lymphovascular invasion in patients treated with radical cystectomy. *BJU Int* 2010; 105: 1402–1412.
17. Als AB, Dyrskjøt L, von der Maase H, et al. Emmprin and survivin predict response and survival following cisplatin-containing chemotherapy in patients with

advanced bladder cancer. *Clin Cancer Res* 2007; 13(15 part 1): 4407–4414.

18. Sanchez-Carbayo M, Socci ND, Lozano J, et al. Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. *J Clin Oncol* 2006; 24: 778–789.
19. Mitra AP, Pagliarulo V, Yang D, et al. Generation of a concise gene panel for outcome prediction in urinary bladder cancer. *J Clin Oncol* 2009; 27: 3929–3937.
20. Karam JA, Lotan Y, Karakiewicz PI, et al. Use of combined apoptosis biomarkers for prediction of bladder cancer recurrence and mortality after radical cystectomy. *Lancet Oncol* 2007; 8: 128–136.
21. Shariat SF, Tokunaga H, Zhou J, et al. p53, p21, pRB, and p16 expression predict clinical outcome in cystectomy with bladder cancer. *J Clin Oncol* 2004; 22: 1014–1024.
22. Michaelson MD, Shipley WU, Heney NM, et al. Selective bladder preservation for muscle-invasive transitional cell carcinoma of the urinary bladder. *Br J Cancer* 2004; 90: 578–581.
23. Sternberg CN, Yagoda A, Scher HI, et al. Preliminary results of M-VAC (methotrexate, vinblastine, doxorubicin and cisplatin) for transitional cell carcinoma of the urothelium. *J Urol* 1985; 133: 403–407.
24. Von der Maase H, Hansen SW, Roberts JT, et al. Gemcitabine and cisplatin versus methotrexate, vinblastine, doxorubicin, and cisplatin in advanced or metastatic bladder cancer: results of a large, randomized, multinational, multicenter, phase III study. *J Clin Oncol* 2000; 18: 3068–3077.
25. Bellmunt J, von der Maase H, Mead GM, et al. Randomized phase III study comparing paclitaxel/cisplatin/gemcitabine and gemcitabine/cisplatin in patients with locally advanced or metastatic urothelial cancer without prior systemic therapy: EORTC30987/Intergroup Study. *Proc Am Soc Clin Oncol* 2007; 25: Abstract 5030.
26. Bellmunt J, von der Maase H, Theodore C, et al. Randomised phase III trial of vinflunine plus best supportive care vs best supportive care alone as 2nd line therapy after a platinum-containing regimen in advanced transitional cell carcinoma of the urothelium. *J Clin Oncol* 2008; 26(suppl 15S): Abstract 5028.
27. Ciardiello F, Caputo R, Bianco R, et al. Antitumor effect and potentiation of cytotoxic drugs activity in human cancer cells by ZD-1839 (Iressa), an epidermal growth factor receptor-selective tyrosine kinase inhibitor. *Clin Cancer Res* 2000; 6: 2053–2063.
28. Petrylak DP, Tangen CM, Van Veldhuizen PJ Jr, et al. Results of the Southwest Oncology Group phase II evaluation (study S0031) of ZD1839 for advanced transitional cell carcinoma of the urothelium. *BJU Int* 2010; 105: 317–321.
29. McConkey DJ, Choi W, Marquis L, et al. Role of epithelial-to-mesenchymal transition (EMT) in drug sensitivity and metastasis in bladder cancer. *Cancer Metastasis Rev* 2009; 28: 335–344.
30. Gallagher DJ, Milowsky MI, Gerst SR, et al. Phase II study of sunitinib in patients with metastatic urothelial cancer. *J Clin Oncol* 2010; 28: 1373–1379.
31. Sonpavde G, Jian W, Liu H, et al. Sunitinib malate is active against human urothelial carcinoma and enhances the activity of cisplatin in a preclinical model. *Urol Oncol* 2009; 27: 391–399.
32. Galsky MD, Sonpavde G, Hellerstedt BA, et al. Phase II study of gemcitabine, cisplatin, and sunitinib in patients with advanced urothelial carcinoma. ASCO Genitourinary Cancers Symposium 2010: Abstract 276.
33. Hahn NM, Stadler WM, Zon RT, et al. A multicenter phase II study of cisplatin (C), gemcitabine (G), and bevacizumab (B) as first-line chemotherapy for metastatic urothelial carcinoma (UC): Hoosier Oncology Group GU-0475. *J Clin Oncol* 2009: 27(suppl 15S): Abstract 5018.
34. Esrig D, Elmajian D, Groshen S, et al. Accumulation of nuclear p53 and tumor progression in bladder cancer. *N Engl J Med* 1994; 331: 1259–1264.
35. Cote RJ, Esrig D, Groshen S, et al. p53 and treatment of bladder cancer. *Nature* 1997; 385: 123–125.
36. Stadler WM, Lerner SP, Groshen S, et al. Randomized trial of p53 targeted adjuvant therapy for patients with organ-confined node-negative urothelial bladder cancer. *J Clin Oncol* 2009; 27(suppl 15S): Abstract 5017.
37. Bellmunt J, Paz-Ares L, Cuello M, et al. Gene expression of ERCC1 as a novel prognostic marker in advanced bladder cancer patients receiving cisplatin-based chemotherapy. *Ann Oncol* 2007; 18: 522–528.
38. Takata R, Katagiri T, Kanehira M, et al. Validation study of the prediction system for clinical response of M-VAC neoadjuvant chemotherapy. *Cancer Sci* 2007; 98: 113–117.
39. Takata R, Katagiri T, Kanehira M, et al. Predicting response to methotrexate, vinblastine, doxorubicin, and cisplatin neoadjuvant chemotherapy for bladder cancers through genome-wide gene expression profiling. *Clin Cancer Res* 2005; 11: 2625–2636.
40. Williams PD, Cheon S, Havaleshko DM, et al. Concordant gene expression signatures predict clinical outcomes of cancer patients undergoing systemic therapy. *Cancer Res* 2009; 69: 8302–8309.
41. Wülfing C, Machiels JP, Richel DJ, et al. A single-arm, multicenter, open-label phase 2 study of lapatinib as the second-line treatment of patients with locally advanced or metastatic transitional cell carcinoma. *Cancer* 2009; 115: 2881–2890.
42. Galsky MD, Zaks T, Hassani H, et al. Target-specific randomized discontinuation trial design: a novel approach in molecular therapeutics. *Invest New Drugs* 2010; 28: 194–198.
43. Galsky MD, Von Hoff DD, Neubauer M, et al. Target-specific, histology-independent, randomized discontinuation study of lapatinib in patients with HER2-amplified solid tumors. *J Clin Oncol* 2009; 27(suppl 15S): Abstract 3541.
44. Pruthi RS, Nielsen M, Heathcote S, et al. A phase II trial of neoadjuvant erlotinib in patients with muscle-invasive bladder cancer undergoing radical cystectomy: clinical and pathological results. *BJU Int* 2010; 106: 349–356.
45. Levitt JM, Yamashita H, Jian W, et al. Dasatinib is preclinically active against Src-overexpressing human transitional cell carcinoma of the urothelium with activated Src signaling. *Mol Cancer Ther* 2010; 9: 1128–1135.
46. Loehrer PJ Sr, Einhorn LH, Elson PJ, et al. A randomized comparison of cisplatin alone or in combination with methotrexate, vinblastine, and doxorubicin in patients with metastatic urothelial

carcinoma: a cooperative group study. *J Clin Oncol* 1992; 10: 1066–1073.

47. Logothetis CJ, Dexeus FH, Finn L, et al. A prospective randomized trial comparing MVAC and CISCA chemotherapy for patients with metastatic urothelial tumors. *J Clin Oncol* 1990; 8: 1050–1055.
48. Siefker-Radtke AO, Millikan RE, Tu SM, et al. Phase III trial of fluorouracil, interferon alpha-2b, and cisplatin versus methotrexate, vinblastine, doxorubicin, and cisplatin in metastatic or unresectable urothelial cancer. *J Clin Oncol* 2002; 20: 1361–1367.
49. Sternberg CN, de Mulder PH, Schornagel JH, et al. Randomized phase III trial of high-dose-intensity methotrexate, vinblastine, doxorubicin, and cisplatin (MVAC) chemotherapy and recombinant human granulocyte colony-stimulating factor versus classic MVAC in advanced urothelial tract tumors: European Organization for Research and Treatment of Cancer Protocol no. 30924. *J Clin Oncol* 2001; 19: 2638–2646.
50. Bamias A, Aravantinos G, Deliveliotis C, et al. Docetaxel and cisplatin with granulocyte colony-stimulating factor (G-CSF) versus MVAC with G-CSF in advanced urothelial carcinoma: a multicenter, randomized, phase III study from the Hellenic Co-operative Oncology Group. *J Clin Oncol* 2004; 22: 220–228.
51. Dreicer R, Manola J, Roth BJ, et al. Phase III trial of methotrexate, vinblastine, doxorubicin, and cisplatin versus carboplatin and paclitaxel in patients with advanced carcinoma of the urothelium. *Cancer* 2004; 100: 1639–1645.
52. Vaughn DJ, Broome CM, Hussain M, et al. Phase II trial of weekly paclitaxel in patients with previously treated advanced urothelial cancer. *J Clin Oncol* 2002; 20: 937–940.
53. McCaffrey JA, Hilton S, Mazumdar M, et al. Phase II trial of docetaxel in patients with advanced or metastatic transitional-cell carcinoma. *J Clin Oncol* 1997; 15: 1853–1857.
54. Witte RS, Elson P, Bono B, et al. Eastern Cooperative Oncology Group phase II trial of ifosfamide in the treatment of previously treated advanced urothelial carcinoma. *J Clin Oncol* 1997; 15: 589–593.
55. Lorusso V, Pollera CF, Antimi M, et al. A phase II study of gemcitabine in patients with transitional cell carcinoma of the urinary tract previously treated with platinum. Italian Co-operative Group on Bladder Cancer. *Eur J Cancer* 1998; 34: 1208–1212.
56. Vaishampayan UN, Faulkner JR, Small EJ, et al. Phase II trial of carboplatin and paclitaxel in cisplatin-pretreated advanced transitional cell carcinoma: a Southwest Oncology Group study. *Cancer* 2005; 104: 1627–1632.
57. Culine S, Theodore C, De Santis M, et al. A phase II study of vinflunine in bladder cancer patients progressing after first-line platinum-containing regimen. *Br J Cancer* 2006; 94: 1395–1401.
58. Galsky MD, Mironov S, Iasonos A, et al. Phase II trial of pemetrexed as second-line therapy in patients with metastatic urothelial carcinoma. *Invest New Drugs* 2007; 25: 265–270.
59. Sweeney CJ, Roth BJ, Kabbinavar FF, et al. Phase II study of pemetrexed for second-line treatment of transitional cell cancer of the urothelium. *J Clin Oncol* 2006; 24: 3451–3457.

Lung Adenocarcinoma: Lessons in Translation from Bench to Bedside

Jaya Sangodkar,[1] Suzanna Katz,[1] Heather Melville,[1] and Goutham Narla, MD, PhD[2]

[1]Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY
[2]Department of Medicine, Mount Sinai School of Medicine, New York, NY

OUTLINE

CASE REPORT
CURRENT TREATMENT MODALITIES
GENETICS OF LUNG CANCER
Epidermal Growth Factor Receptor (EGFR) Signaling Pathway
TARGETING LUNG CANCER
Cetuximab and Panitumumab
Erlotinib and Gefitinib
Resistance to Erlotinib Response via EGFR Somatic Mutation
FUTURE OF TARGETED MOLECULAR THERAPY

ABSTRACT

Lung cancer is currently the leading cause of cancer-related death in the United States and worldwide, accounting for approximately a third of all cancer diagnoses and deaths. The American Cancer Society estimates 159,390 deaths from lung cancer in the United States for 2009 alone. The high mortality rate associated with lung cancer has prompted numerous exhaustive efforts to identify novel therapeutic targets and treatment modalities for this deadly disease. The characterization of signaling pathways underlying the development and progression of lung cancer, particularly lung adenocarcinoma, has been instrumental in developing novel therapeutic strategies to target aggressive metastatic disease. This paradigm is best illustrated in lung adenocarcinoma by recent studies of the epidermal growth factor receptor, which has been identified to be a critical diagnostic and therapeutic target. Our understanding of the central role of epidermal growth factor receptor in the development and progression of lung adenocarcinoma has led to the development of molecular agents against this key oncogene that have demonstrated significant clinical efficacy against the disease. Despite these successes, de novo or acquired resistance to these anti–epidermal growth factor receptor agents invariably develops, either through additional mutations in the epidermal growth factor receptor, or abnormal regulation of downstream signaling pathways. Thus, it is necessary to further investigate the molecular determinants of treatment resistance and to develop novel therapeutic strategies directed specifically against the molecular drivers of metastatic chemoresistant lung cancer. Given the heterogeneity and convergence of signaling pathways underlying both disease etiology and chemoresistance, future efforts to administer rationally designed agents against multiple molecular targets could serve to improve both first-line and second-line therapies for patients with lung adenocarcinoma. ***Mt Sinai J Med 77:597–605, 2010.*** *© 2010 Mount Sinai School of Medicine*

Key Words: genomics, lung cancer, targeted molecular therapy.

Address Correspondence to:

Goutham Narla
Mount Sinai School of Medicine
New York, NY
Email: goutham.narla@mssm.edu

Lung cancer is currently the leading cause of cancer-related death in the United States and worldwide, accounting for approximately a third of all cancer diagnoses and deaths. The American Cancer Society estimates 159,390 deaths from lung cancer in the United States for 2009 alone.[1] Lung cancer can be broadly divided into 2 major histological categories: small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC). Small-cell lung cancer comprises about 20% of all diagnosed cases, whereas NSCLC accounts for the remaining 80% of cases. Non-small-cell lung cancer can be subdivided into the following major histological subtypes: squamous cell carcinoma, large cell carcinoma, and adenocarcinoma.[2]

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20226

© 2010 Mount Sinai School of Medicine

Lung adenocarcinoma is the most common subtype of NSCLC in the United States.[3]

CASE REPORT

A 50-year-old nonsmoker male presented with progressive fatigue and weight loss. He was otherwise healthy with no significant past medical history, and no family history of malignancy. Computed tomography scans of the chest and brain revealed a large pulmonary mass (4 cm × 4.6 cm) on the left lower lobe, enlarged contralateral mediastinal lymphadenopathy, and several hypoattenuating lesions of varying sizes in the liver. A diagnosis of metastatic lung adenocarcinoma was made, and cytotoxic chemotherapy with paclitaxel and carboplatin was started. Despite an initial response to therapy, the tumor progressed and second-line chemotherapy was started unsuccessfully in this patient. After subsequent failure on salvage cytotoxic chemotherapy, erlotinib therapy was administered and a significant partial response was achieved. The lung mass nearly disappeared and the lesions in the liver decreased on average by >50% in volume after initiation of anti–epidermal growth factor receptor (EGFR) therapy. Subsequent sequence analysis of the initial tumor biopsy revealed a 15-nucleotide heterozygous mutation in exon 19 of the EGFR. Seven months after initiation of erlotinib monotherapy, the disease progressed and the patient ultimately died of recurrent chemoresistant metastatic lung adenocarcinoma.

CURRENT TREATMENT MODALITIES

After a NSCLC tumor has been properly staged, a treatment modality is chosen based upon a number of factors including age, comorbidity, and performance status of the patient. In general, stage I and II tumors are surgically resectable and approached with curative intent. Surgery is the first line of therapy and treatment of choice for patients presenting these tumors,[4–6] and the standard of care involves a lobectomy[6,7] or a limited (sublobar) resection in some instances.[6,7] Video-assisted thoracoscopic surgery is a minimally invasive procedure used for lobectomies and limited resection surgeries in selected patients[6,8,9] whereas en bloc resection has been used for tumors with chest wall involvement.[6,10] Adjuvant chemotherapy is sometimes useful for lowering the risk of distant metastases after surgical resection of a tumor,[6,11,12] showing a small survival advantage in stage IB disease with T2 tumors >4 cm.[1,6,12,13]

Most patients with NSCLC, however, will be diagnosed with advanced disease, and their overall survival rate remains dismal, with a 5-year survival rate of 8%–10%.[14–16] As in the case of stage I and II disease, surgery is still an option for a small number of patients with stage III disease (and those with separate tumor nodules other than the primary tumor); for these patients, adjuvant chemotherapy or adjuvant chemoradiotherapy is administered following surgery.[16–18] Most patients with stage IV NSCLC will undergo platinum-based chemotherapy (cisplatin or carboplatin),[19–21] either as a combined modality or as a primary palliative treatment for malignant effusions.[16,17] Some patients with advanced disease will be eligible for non–platinum-based combination; for example, novel tyrosine kinase inhibitors (TKIs) erlotinib or gefitinib,[19,22] which have demonstrated variable clinical efficacy. Radiation therapy (RT) is the standard of care for those not undergoing combined modality treatments, with definitive RT administered to patients with unresectable NSCLC[16,23] in doses that are made more tolerable with the use of amifostine to protect against RT-induced toxicity.[16,24–28] Stage IV patients are approached with palliative systemic therapy; these patients may present either with disseminated metastases, a malignant pleural or pericardial effusion, or with relapsed disease after having received definitive treatment.[19]

Because many patients do not respond to first-line therapy, there are currently 3 second-line treatment options approved by the US Food and Drug Administration (FDA) for advanced NSCLC: docetaxel and pemetrexed (which are cytotoxic chemotherapy agents), as well as erlotinib (a TKI).[29–34] Docetaxel was the first drug to be approved by the FDA as a second-line therapy, wereas erlotinib is currently FDA approved as a second-line therapy and even third-line therapy in NSCLC; on the other hand, based on clinical trial results, the FDA restricted the use of pemetrexed for second-line therapy in NSCLC of only the nonsquamous cell histology.[30,31,35,36]

In summary, evidence from large randomized clinical trials suggests that although early-stage lung cancer is treatable, the majority of cases are diagnosed in more advanced or metastatic stages, which leads to poor outcomes and dismal overall survival.

A better and more thorough understanding of the molecular determinants of lung cancer progression is imperative for the development of rational targeted strategies for the treatment of this disease.

Based upon these observations, it is clear that a better and more thorough understanding of the molecular determinants of disease progression is imperative for the development of rational targeted strategies for the treatment of lung cancer.

GENETICS OF LUNG CANCER

Genetic alterations that have been identified in the development and progression of lung cancer can be grouped into 2 classes: loss of tumor-suppressor gene function, or oncogene activation. Tumor-suppressor genes that are frequently inactivated in lung cancer include p53, Rb, phosphatase and tensin homolog (PTEN), and p16. This inactivation occurs through loss of heterozygosity, somatic mutation, decreased expression, and/or promoter hypermethylation. Along with the loss of tumor-suppressor gene function, oncogene activation is a hallmark of lung adenocarcinoma, with the key molecular alterations occurring in the following proto-oncogenes: c-Myc, Cyclin D1, BCL-2, K-Ras, and EGFR.[3] Of these specific genetic alterations, EGFR activation has become an important pathway of study as both a key molecular driver of disease progression as well as a target for the development of molecular agents to treat disseminated disease. In lung adenocarcinoma, EGFR activation has been shown to occur through receptor amplification, somatic mutation, and/or downstream activation of the EGFR-depending signaling pathways.[37–40] Retrospective analysis of samples derived from lung adenocarcinoma patients has demonstrated that EGFR is overexpressed in >60% of NSCLC cases. Furthermore, this overexpression is associated with poor prognosis.[41] At the molecular level, activation of EGFR signaling has a diverse range of cellular effects, including increased proliferation, metastasis, and angiogenesis, and decreased apoptosis.[42] Combined, these findings make the EGFR and its related family members prime candidates for the development of targeted therapeutics to treat lung cancer.

Epidermal Growth Factor Receptor (EGFR) Signaling Pathway

The EGFR is a 170-kDa transmembrane glycoprotein encoded by 28 exons that span about 190,000 nucleotides on chromosome 7p12.[43] As a member of the tyrosine kinase family of receptors, the EGFR belongs to the subfamily of 4 closely related receptors: HER-1/ErbB1, HER-2/neu/ErbB2, HER-3/ErbB3, and HER-4/ErbB4.[44] Structurally, the EGFR comprises an extracellular ligand-binding domain and an intracellular protein tyrosine kinase domain connected by a small transmembrane-anchoring region (Figure 1). Signaling of EGFR is triggered by the binding of growth factors, such as epidermal growth factor (EGF), transforming growth factor-alpha, or amphiregulin.[43] Ligand binding results in homodimerization of the receptor, or heterodimerization with another closely related receptor such as HER-2/neu/ErbB2. Activation of the receptor dimer leads to downstream signaling cascades including the Ras/mitogen-activated protein kinase and phosphoinositide 3-kinase (PI3K) signaling pathways, promoting increased cellular proliferation and survival (Figure 2).

Although EGFR is the primary target in designing and developing efficacious therapies for NSCLC, another critical signaling ligand, vascular endothelial growth factor (VEGF), has been implicated in the development and progression of the disease, playing a significant role in promoting angiogenesis

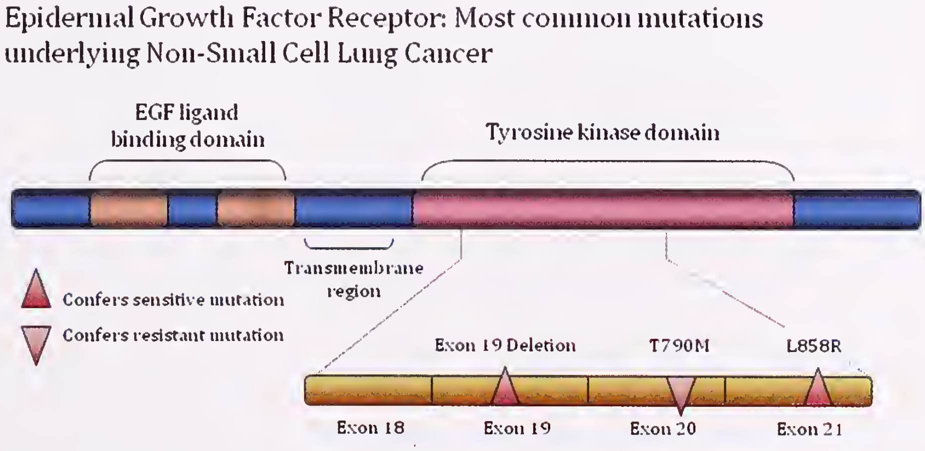

Epidermal Growth Factor Receptor: Most common mutations underlying Non-Small Cell Lung Cancer
EGF ligand binding domain
Tyrosine kinase domain
Transmembrane region
Confers sensitive mutation
Confers resistant mutation
Exon 19 Deletion
T790M
L858R
Exon 18
Exon 19
Exon 20
Exon 21

Fig 1. EGFR: Most common mutations underlying NSCLC. **Abbreviations:** EGFR, epidermal growth factor receptor; NSCLC, non-small-cell lung cancer.

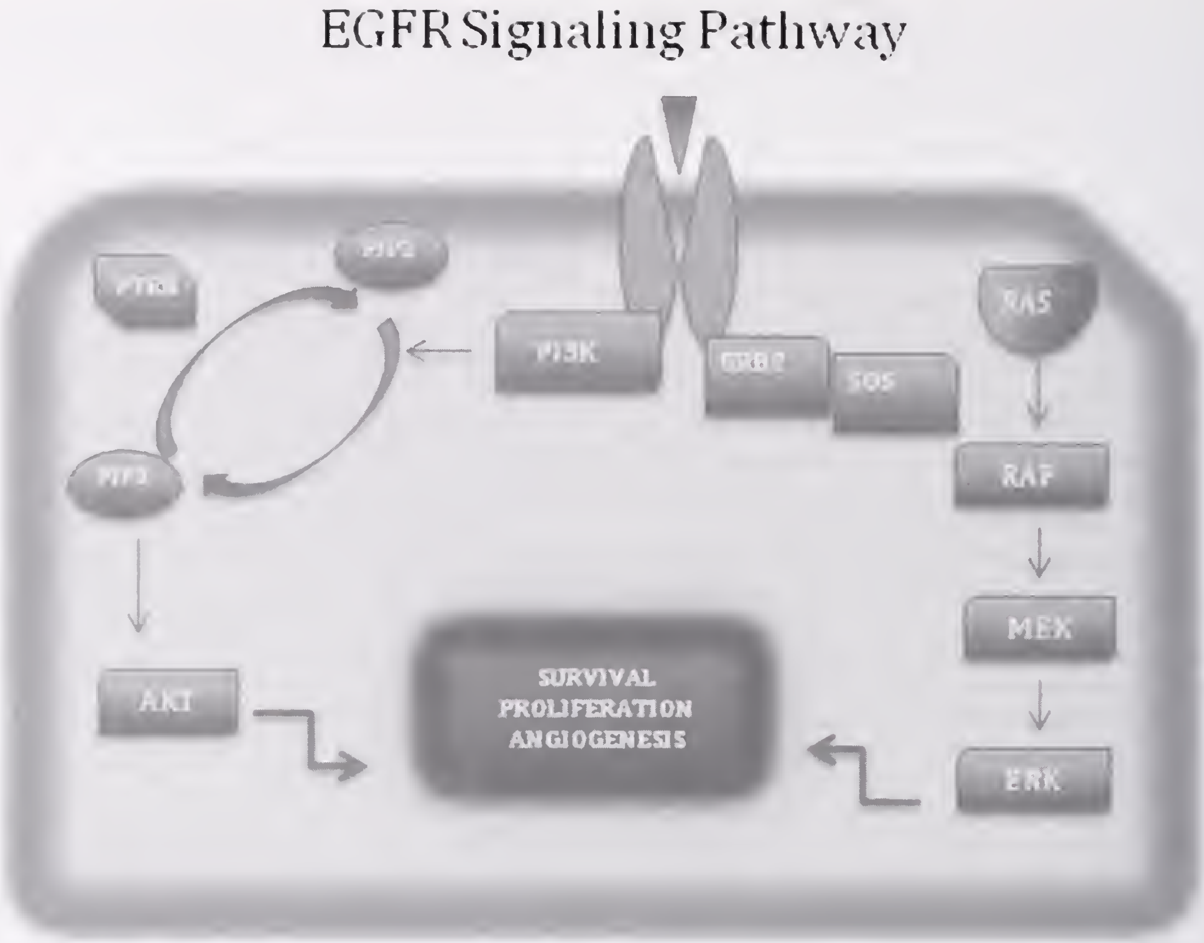


Fig 2. EGFR signaling pathway. **Abbreviations:** EGFR, epidermal growth factor receptor; ERK, extracellular signal-regulated kinase; GRB2, growth factor receptor bound protein-2; MEK, mitogen-activated protein kinase and extracellular signal-regulated kinase; PIP2, phosphatidylinositol 4,5-bisphosphate; PIP3, phosphatidylinositol (3,4,5)-trisphosphate; PTEN, phosphatase and tensin homolog; SOS, son of sevenless. This figure is available in color online at Wiley Online Library (wileyonlinelibrary.com).

Table 1. *Lung Cancer Pathway Targets.*

Pathway Target	Drug Type	Drug Name	Current Trial Status for NSCLC
EGFR	Small-molecule TKI	Erlotinib	FDA approved
		Gefinitib	Phase III
	Monoclonal antibody	Cetuximab	Phase III, II, I
		Panitumumab	Phase I, II
VEGF, VEGFR	Monoclonal antibody	Bevacizumab	Phase III
	Small-molecule TKI	Sunitinib	Phase II
		Cediranib	Phase II
		Sorafenib	Phase II, III
EGFR, VEGFR	Small-molecule TKI	Vandetanib	Phase II, III

Abbreviations: EGFR, epidermal growth factor receptor; FDA, Food and Drug Administration; NSCLC, non-small-cell lung cancer; TKI, tyrosine kinase inhibitor; VEGF, vascular endothelial growth factor; VEGFR, vascular endothelial growth factor receptor.

and metastatic potential of cancer cells.[45] Combined inhibition of both VEGF and EGFR has been explored in ongoing clinical trials with agents designed to abrogate these aberrant signaling pathways in lung adenocarcinoma.[46] The monoclonal antibody bevacizumab and small-molecule TKIs of VEGF, such as sunitinib, cediranib, and sorafenib, are currently being tested in phase I, II, and III trials with applications for FDA approval in submission from major pharmaceutical companies (Table 1). Because of the potent oncogenic effects and importance in lung cancer development and progression, the EGFR and VEGF signaling pathways have become major targets for the treatment of metastatic lung cancer.[47–49]

TARGETING LUNG CANCER

Several targeted molecular therapies have been developed to target the aberrant EGFR signaling pathway that drive the development and progression of lung adenocarcinoma. These therapies include anti-EGFR monoclonal antibodies, such as cetuximab and panitumumab, and small-molecule inhibitors, such as erlotinib and gefitinib. Of these therapies, only erlotinib is currently approved by the FDA for use in patients with metastatic NSCLC,[50] whereas others are still in various stages of clinical trial (Table 1).

Cetuximab and Panitumumab

Monoclonal antibodies (mAbs), such as cetuximab and panitumumab, have been designed to target and inhibit EGFR (Table 1). Cetuximab has demonstrated clinical efficacy for a variety of cancers and is currently FDA-approved for treatment of metastatic colorectal cancer and squamous cell carcinoma of the head and neck; however, further studies are under way to determine its efficacy in NSCLC. Early clinical trials have shown a role for cetuximab, a human-mouse chimeric mAb, in the treatment of NSCLC,[51] particularly the recent clinical phase III First-Line Erbitux in Lung Cancer (FLEX) trial.[52] In this study, >1100 NSCLC patients with positive EGFR tumor status confirmed by histological analysis of the tumor were treated with combination cetuximab/chemotherapy or chemotherapy alone. Kaplan-Meier survival analysis indicated that adjuvant treatment with cetuximab prolongs lifespan in an EGFR-selected population of patients, providing evidence that further investigation of those patients who respond to cetuximab may reveal the molecular dynamics that confer sensitivity to this mAb therapy. Interestingly, a more recent study that did not restrict to specific EGFR-mutational status showed a minimal and insignificant survival benefit to adjuvant cetuximab therapy in combination with traditional chemotherapy, thereby casting some doubt on the efficacy of these mAbs in treating NSCLC in the general, unselected population of lung adenocarcinoma patients.[53]

Panitumumab, a fully human monoclonal EGFR antibody, is undergoing continued clinical investigation of safety and efficacy for lung adenocarcinoma.[54] Early clinical trials were unfortunately disappointing, showing no significant extension of progression-free disease when treatment regimen was supplemented with panitumumab.[55] Further studies with this mAb in combination with traditional chemotherapies and a small-molecule inhibitor, AMG 706 (Amgen Corporation), have been initiated with the hope that targeting multiple aspects of aberrant signaling will more effectively combat aggressive tumors.[56]

Erlotinib and Gefitinib

Erlotinib and gefitinib are small molecules designed to specifically target EGFR tyrosine kinase activity by binding to the adenosine triphosphate binding cassette of the receptor, therefore inhibiting EGFR autophosphorylation and abrogating aberrant downstream signaling.[57] Several clinical studies have demonstrated antitumor activity for erlotinib in patients who either failed to respond to traditional cytotoxic chemotherapy or who developed recurrent disease on first-line therapy. Furthermore, in a large, double-blind, placebo-controlled clinical study of >800 patients with chemotherapy-resistant metastatic lung cancer, erlotinib was shown to increase overall survival by approximately 2 months. Based upon these results, the FDA approved the use of erlotinib in patients with treatment-resistant metastatic chemoresistant disease, and more recently for maintenance treatment after traditional chemotherapy.[43,58] Although the improvement in median survival was modest at best, there were specific subgroups of patients who demonstrated dramatic responses to erlotinib monotherapy. This subset of patients included nonsmokers, women, and people of Asian descent.[2] Subsequent molecular studies of the EGFR and downstream effectors of this signaling pathway revealed that treatment responders harbored activating mutations in the EGFR, whereas those who failed to respond to anti–EGFR-based therapy had alterations in downstream signaling mediators of the EGFR signaling pathway including Ras, PI3K, and PTEN.[59–64] In a recent study, a subset of these EGFR-mutant NSCLC patients were selected for and treated with gefinitib, another small-molecule tyrosine kinase inhibitor that demonstrates similar activity to erlotinib. This clinical trial, the Iressa Pan-Asia Study (IPASS), utilized geftinib as a first-line monotherapy for otherwise untreated patients in comparison with traditional chemotherapy, carboplatin-paclitaxel. Although gefitinib is not currently FDA-approved for use in the United States, the results indicate that within a selected subgroup of EGFR-mutant lung adenocarcinoma patients, gefinitib significantly extended progression-free survival (mean 10.8 months, versus 5.4 months for chemotherapy) and caused fewer adverse side effects, therefore providing evidence for the use of gefitinib as a first-line monotherapy for patients with EGFR-mutant lung cancer.[65] This data suggests that patients with specific molecular alteration mutations that result in activated oncogenic EGFR signaling may be spared the side effects of traditional cytotoxic chemotherapy through the use of targeted molecular therapies such as gefitinib or erlotinib. On the other hand, a negative EGFR status should also indicate treatment with gefitinib will unlikely yield a positive response, and therefore these patients should receive traditional chemotherapy. Screening and determination of each individual tumor genotype is therefore critical in predicting response and success with these targeted therapies and has laid the groundwork for use of personalized medicine–based approaches for the treatment of lung cancer.

> *Screening and determination of each individual tumor genotype is critical in predicting response and success with targeted therapies and has laid the groundwork for use of personalized medicine–based approaches for the treatment of lung cancer.*

Resistance to Erlotinib Response via EGFR Somatic Mutation

Although the IPASS study selected for EGFR mutation and demonstrated gefitinib response in these patients, analysis of unselected NSCLC samples has shown that only 10% of all tumors derived from patients in North America and Western Europe possessed EGFR mutations, whereas the occurrence of these same somatic mutations in individuals of East Asian descent is 30%–50%.[2] More than 80% of these mutations were either L858R point mutations within exon 21 of the EGFR gene or in-frame deletions within exon 19 (Figure 1). Retrospective analysis of erlotinib-treated patient cohorts revealed that tumors harboring these activating mutations responded significantly better to erlotinib compared with patients with other oncogenic mutations. Further molecular evidence to strengthen this finding was provided through the use of mouse models of the disease. Specifically, an in vivo transgenic mouse model was designed to study the L858R mutant and the exon 19 deletion mutant in type II pneumocytes under the control of doxycycline. These transgenic mice developed lung adenocarcinoma with 100% penetrance. Interestingly, inhibition of EGFR signaling in vivo, either through withdrawal of doxycycline or treatment with erlotinib, resulted in marked tumor regression. Combined, these experiments demonstrate that activating EGFR mutations is necessary and sufficient to drive the development and progression of lung cancer, and that inhibition of EGFR signaling with erlotinib induces apoptosis and tumor regression in vivo.[66] Yet, not all EGFR kinase mutations are associated with hypersensitivity to erlotinib. For example, the T790M mutation (Table 1) in the EGFR alters the binding kinetics of erlotinib to the EGFR, thus rendering tumor cells resistant to erlotinib.[67] Interestingly, the T790M mutation is associated with both acquired resistance and primary resistance to erlotinib in human lung adenocarcinoma.[50] Recent studies suggest that up to 50% of erlotinib-resistant tumors harbor the "gatekeeper" T790M EGFR mutation.[68] Additional mechanisms of resistance to EGFR inhibitors include hyperactivation of downstream signaling components that allow for circumvention of EGFR inhibition. Examples of this resistance mechanism include activating mutations in the proto-oncogene Ras, which result in increased downstream signaling through the mitogen-activated protein kinase/extracellular signal-regulated kinase signaling pathway, thereby promoting cellular proliferation and survival. In fact, recent evidence from several clinical trials suggests that screening for activating RAS mutations is critical for determining patient response to anti–EGFR-based therapies. In addition, activated AKT signaling can occur through several mechanisms, such as PTEN depletion[69] and PI3K activating mutations,[70] and is associated with either de novo or acquired resistance to anti–EGFR-based therapies.

FUTURE OF TARGETED MOLECULAR THERAPY

Understanding the molecular determinants driving the development and progression of lung adenocarcinoma is clearly a critical area of research inquiry, and will allow for the development of molecular therapies targeted to the specific alterations that drive treatment resistance and failure in any given patient tumor. Furthermore, screening for mutations that may either induce hypersensitivity or resistance to any given drug could be used for more appropriate patient stratification prior to treatment initiation. For example, recent evidence from several randomized studies has demonstrated that patients harboring EGFR mutations should be offered erlotinib as first-line therapy rather than as second-line therapy after treatment with traditional cytotoxic chemotherapy has failed.

> *Recent evidence from several randomized studies has demonstrated that patients harboring epidermal growth factor receptor mutations should be offered erlotinib as first-line therapy rather than as second-line therapy after treatment with traditional cytotoxic chemotherapy has failed.*

Regardless of stage, patients harboring downstream mutations in PI3K, Ras, or PTEN should not be given EGFR-based therapies at any point of their treatment regimen. This thorough understanding of the downstream regulators of EGFR signaling and the molecular mechanisms involved in activation of this signaling pathway has begun to lay the foundation for the development of personalized and targeted approaches for the treatment of patients suffering from lung adenocarcinoma. As is the case with traditional cytotoxic chemotherapies or newer targeted molecular agents, resistance to anti–EGFR-based therapies invariably develops. Future research, therefore, must focus on identifying additional molecular alterations and signaling pathways to help guide the decisions for rationally designed therapies and to allow for the development of new classes of targeted molecular agents for this incurable disease.

> *As is the case with traditional cytotoxic chemotherapies or newer targeted molecular agents, resistance to anti–EGFR-based therapies invariably develops. Future research, therefore, must focus on identifying additional molecular alterations and signaling pathways to help guide the decisions for rationally designed therapies and to allow for the development of new classes of targeted molecular agents for this incurable disease.*

An approach to restore sensitivity to cell lines with activated downstream EGFR signaling would involve the understanding of mechanisms to reactivate the apoptotic process and restore these downstream signaling events in vivo. Through an active and dynamic collaboration between the basic sciences, coupled with the clinical application of these findings to patients suffering from lung adenocarcinoma, we have, for the first time in recent years, begun to make significant advances in the treatment of this still deadly and lethal disease.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Spira A, Ettinger DS. Drug therapy: multidisciplinary management of lung cancer. *N Engl J Med* 2004; 350: 379–392.
2. Sharma SV, Bell DW, Settleman J, et al. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 2007; 7: 169–181.
3. Minna JD, Roth JA, Gazdar AF. Focus on lung cancer. *Cancer Cell* 2002; 1: 49–52.
4. Raz DJ, Zell JA, Ou SHI, et al. Natural history of stage I non-small cell lung cancer: implications for early detection. *Chest* 2007; 132: 193–199.
5. Manser R, Wright G, Hart D, et al. Surgery for local and locally advanced non-small cell lung cancer. *Cochrane Database Syst Rev.* 2005; 1: CD004699.
6. Thomas KW, Gould MK. Diagnosis and staging of non-small cell lung cancer. In: Basow DS, ed. *UpToDate.* Waltham, MA: UpToDate; 2010.
7. Holmes CE, Ruckdeschel JC, Johnston M, et al. Randomized trial of lobectomy versus limited resection for T1 N0 non-small-cell lung cancer. *Ann Thorac Surg* 1995; 60: 615–622.
8. Yan TD, Black D, Bannon PG, et al. Systematic review and meta-analysis of randomized and nonrandomized trials on safety and efficacy of video-assisted thoracic surgery lobectomy for early-stage non-small-cell lung cancer. *J Clin Oncol* 2009; 27: 2553–2562.
9. Mentzer SJ, DeCamp MM, Harpole DH Jr, et al. Thoracoscopy and video-assisted thoracic surgery in the treatment of lung cancer. *Chest* 1995; 107(6 suppl): 298S–301S.
10. Doddoli C, D'Journo B, Le Pimpec-Barthes F, et al. Lung cancer invading the chest wall: a plea for en-bloc resection but the need for new treatment strategies. *Ann Thorac Surg* 2005; 80: 2032–2040.
11. Niiranen A, Niitamokorhonen S, Kouri M, et al. Adjuvant chemotherapy after radical surgery for non-small-cell lung cancer: a randomized study. *J Clin Oncol* 1992; 10: 1927–1932.

12. Hanna N. Adjuvant systemic therapy in resectable non-small cell lung cancer. In: Basow DS, ed. *UpToDate*. Waltham, MA: UpToDate; 2010.
13. Scott WJ, Howington J, Feigenberg S, et al. Treatment of non-small cell lung cancer stage I and stage II: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest* 2007; 132: 234S–242S.
14. Beer DG, Kardia SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002; 8: 816–824.
15. Fry WA, Phillips JL, Menck HR. Ten-year survey of lung cancer treatment and survival in hospitals in the United States: a national cancer data base report. *Cancer* 1999; 86: 1867–1876.
16. Schild SE, Ramalingam SS. Management of stage III non-small cell lung cancer. In: Basow DS, ed. *UpToDate*. Waltham, MA: UpToDate; 2010.
17. Osaki T, Sugio K, Hanagiri T, et al. Survival and prognostic factors of surgically resected T4 non-small cell lung cancer. *Ann Thorac Surg* 2003; 75: 1745–1751.
18. Port JL, Korst RJ, Lee PC, et al. Surgical resection for multifocal (T4) non-small cell lung cancer: is the T4 designation valid? *Ann Thorac Surg* 2007; 83: 397–400.
19. Lilenbaum R, Samuels M, Taffaro-Neskey M, et al. Phase II trial of combined modality therapy with myeloid growth factor support in patients with locally advanced non-small cell lung cancer. *J Thorac Oncol* 2010; 5: 837–840.
20. Bunn PA Jr. Clinical experiences with carboplatin (paraplatin) in lung cancer. *Semin Oncol* 1992; 19(1 suppl 2): 1–11.
21. Bunn PA. The expanding role of cisplatin in the treatment of non-small-cell lung cancer. *Semin Oncol* 1989; 16: 10–21.
22. Giaccone G, Gallegos Ruiz M, Le Chevalier T, et al. Erlotinib for frontline treatment of advanced non-small cell lung cancer: a phase II study. *Clin Cancer Res* 2006; 12(20 part 1): 6049–6055.
23. Roswit B, Patno ME, Rapp R, et al. The survival of patients with inoperable lung cancer: a large-scale randomized study of radiation therapy versus placebo. *Radiology* 1968; 90: 688–697.
24. Antonadou D, Throuvalas N, Petridis A, et al. Effect of amifostine on toxicities associated with radiochemotherapy in patients with locally advanced non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys* 2003; 57: 402–408.
25. Leong SS, Tan EH, Fong KW, et al. Randomized double-blind trial of combined modality treatment with or without amifostine in unresectable stage III non-small-cell lung cancer. *J Clin Oncol* 2003; 21: 1767–1774.
26. Senzer N. A phase III randomized evaluation of amifostine in stage IIIA/IIIB non-small cell lung cancer patients receiving concurrent carboplatin, paclitaxel, and radiation therapy followed by gemcitabine and cisplatin intensification: preliminary findings. *Semin Oncol* 2002; 29(6 suppl 19): 38–41.
27. Antonadou D, Coliarakis N, Synodinou M, et al. Randomized phase III trial of radiation treatment +/− amifostine in patients with advanced-stage lung cancer. *Int J Radiat Oncol Biol Phys* 2001; 51: 915–922.
28. Koukourakis MI, Kyrias G, Kakolyris S, et al. Subcutaneous administration of amifostine during fractionated radiotherapy: a randomized phase II study. *J Clin Oncol* 2000; 18: 2226–2233.
29. Shepherd FA, Pereira JR, Ciuleanu T, et al. Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* 2005; 353: 123–132.
30. Shepherd FA, Dancey J, Ramlau R, et al. Prospective randomized trial of docetaxel versus best supportive care in patients with non-small-cell lung cancer previously treated with platinum-based chemotherapy. *J Clin Oncol* 2000; 18: 2095–2103.
31. Hanna N, Shepherd FA, Fossella FV, et al. Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-cell lung cancer previously treated with chemotherapy. *J Clin Oncol* 2004; 22: 1589–1597.
32. Fossella FV, DeVore R, Kerr RN, et al. Randomized phase III trial of docetaxel versus vinorelbine or ifosfamide in patients with advanced non-small-cell lung cancer previously treated with platinum-containing chemotherapy regimens. *J Clin Oncol* 2000; 18: 2354–2362.
33. Triano LR, Deshpande H, Gettinger SN. Management of patients with advanced non-small cell lung cancer: current and emerging options. *Drugs* 2010; 70: 167–179.
34. Sheth S. Current and emerging therapies for patients with advanced non-small-cell lung cancer. *Am J Health Syst Pharm* 2010; 67(1 suppl 1): S9–S14.
35. Scagliotti GV, Parikh P, von Pawel J, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. *J Clin Oncol* 2008; 26: 3543–3551.
36. Peterson P, Park K, Fossella FV, et al. Is pemetrexed more effective in adenocarcinoma and large cell lung cancer than in squamous cell carcinoma? A retrospective analysis of a phase III trial of pemetrexed vs docetaxel in previously treated patients with advanced non-small cell lung cancer (NSCLC). *J Thorac Oncol* 2007; 2(suppl 4): S851.
37. Farray D, Mirkovic N, Albain KS. Multimodality therapy for stage III non-small-cell lung cancer. *J Clin Oncol* 2005; 23: 3257–3269.
38. Laurie NA, Donovan SL, Shih CS, et al. Inactivation of the p53 pathway in retinoblastoma. *Nature* 2006; 444: 61–66.
39. Knudsen ES, Knudsen KE. Tailoring to RB: tumour suppressor status and therapeutic response. *Nat Rev Cancer* 2008; 8: 714–724.
40. Croce CM. Molecular origins of cancer: oncogenes and cancer. *N Engl J Med* 2008; 358: 502–511.
41. Olsen JV, Blagoev B, Gnad F, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006; 127: 635–648.
42. Rikova K, Guo A, Zeng Q, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 2007; 131: 1190–1203.
43. Pao W, Miller VA. Epidermal growth factor receptor mutations, small-molecule kinase inhibitors, and non-small-cell lung cancer: current knowledge and future directions. *J Clin Oncol* 2005; 23: 2556–2568.
44. Oda K, Matsuoka Y, Funahashi A, et al. A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol Syst Biol* 2005; doi:10.1038/msb4100014.
45. Korpanty G, Smyth E, Sullivan LA, et al. Antiangiogenic therapy in lung cancer: focus on vascular endothelial growth factor pathway. *Exp Biol Med* 2010; 235: 3–9.

46. Natale RB. Dual targeting of the vascular endothelial growth factor receptor and epidermal growth factor receptor pathways with vandetinib (ZD6474) in patients with advanced or metastatic non-small cell lung cancer. *J Thorac Oncol* 2008; 3(6 suppl 2): S128–S130.
47. Herbst RS. Review of epidermal growth factor receptor biology. *Int J Radiat Oncol Biol Phys* 2004; 59: 21–26.
48. Downward J. Targeting ras signalling pathways in cancer therapy. *Nat Rev Cancer* 2003; 3: 11–22.
49. Hennessy BT, Smith DL, Ram PT, et al. Exploiting the PI3K/AKT pathway for cancer drug discovery. *Nat Rev Drug Discov* 2005; 4: 988–1004.
50. Ciardiello F, Tortora G. Drug therapy: EGFR antagonists in cancer treatment. *N Engl J Med* 2008; 358: 1160–1174.
51. Joy AA, Butts CA. Extending outcomes: epidermal growth factor receptor–targeted monoclonal antibodies in non-small-cell lung cancer. *Clin Lung Cancer* 2009; 10(suppl 1): S24–S29.
52. Pirker R, Pereira JR, Szczesna A, et al. Cetuximab plus chemotherapy in patients with advanced non-small-cell lung cancer (FLEX): an open-label randomised phase III trial. *Lancet* 2009; 373: 1525–1531.
53. Lynch TJ, Patel T, Dreisbach L, et al. Cetuximab and first-line taxane/carboplatin chemotherapy in advanced non-small-cell lung cancer: results of the randomized multicenter phase III trial BMS099. *J Clin Oncol* 2010; 28: 911–917.
54. Socinski MA. Antibodies to the epidermal growth factor receptor in non small cell lung cancer: current status of matuzumab and panitumumab. *Clin Cancer Res* 2007; 13(15 part 2): S4597–S4601.
55. Crawford J, Sandler AB, Hammond LA, et al. ABX-EGF in combination with paclitaxel and carboplatin for advanced non-small cell lung cancer (NSCLC). *J Clin Oncol* 2004; 22(suppl 14S): abstract 7083.
56. Blumenschein G Jr, Sandler A, O'Rourke T, et al. Safety and pharmacokinetics of AMG 706, panitumumab, and carboplatin/paclitaxel for the treatment of patients with advanced non-small cell lung cancer. *J Clin Oncol* 2006; 24(suppl 18S): abstract 7119.
57. Kumar B, Cordell KG, Lee JS, et al. Clinical implications of EGFR expression, HPV titer, and smoking status in advanced stage oropharyngeal squamous cell carcinoma patients. *Int J Biol Markers* 2007; 22: 67.
58. Tsao MS, Sakurada A, Cutz JC, et al. Erlotinib in lung cancer: molecular and clinical predictors of outcome. *N Engl J Med* 2005; 353: 133–144.
59. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004; 304: 1497–1500.
60. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004; 350: 2129–2139.
61. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci U S A*. 2004; 101: 13306–13311.
62. Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *Int J Cancer* 2006; 118: 257–262.
63. Zhou S, Ren S, Yan L, et al. Clinical efficacy of erlotinib in patients previously treated for advanced non-small cell lung cancer. *Respirology* 2009; 14: 709–715.
64. Kosaka T, Yatabe Y, Endoh H, et al. Mutations of the epidermal growth factor receptor gene in lung cancer: biological and clinical implications. *Cancer Res* 2004; 64: 8919–8923.
65. Maemondo M, Inoue A, Kobayashi K, et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. *N Engl J Med* 2010; 362: 2380–2388.
66. Politi K, Zakowski MF, Fan PD, et al. Lung adenocarcinomas induced in mice by mutant EGF receptors found in human lung cancers respond to a tyrosine kinase inhibitor or to down-regulation of the receptors. *Genes Dev* 2006; 20: 1496–1510.
67. Yun CH, Mengwasser KE, Toms AV, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci U S A* 2008; 105: 2070–2075.
68. Politi K, Fan PD, Shen R, et al. Erlotinib resistance in mouse models of epidermal growth factor receptor–induced lung adenocarcinoma. *Dis Model Mech* 2010; 3: 111–119.
69. Sos ML, Koker M, Weir BA, et al. PTEN loss contributes to erlotinib resistance in EGFR-mutant lung cancer by activation of Akt and EGFR. *Cancer Res* 2009; 69: 3256–3261.
70. La Monica S, Galetti M, Alfieri RR, et al. Everolimus restores gefitinib sensitivity in resistant non-small cell lung cancer cell lines. *Biochem Pharmacol* 2009; 78: 460–468.

Advanced or Metastatic Pancreatic Cancer: Molecular Targeted Therapies

Soley Bayraktar, MD and Caio M. Rocha-Lima, MD

Sylvester Comprehensive Cancer Center, Division of Hematology/Oncology,
University of Miami, Miami, FL

OUTLINE

KIRSTEN RAT SARCOMA ONCOGENE INHIBITORS
  Matrix Metalloproteinases Inhibitors
  Cyclooxygenase-2 Inhibitors
  PI3K, Akt, and mTOR Inhibitors
THERAPIES DIRECTED AGAINST EPIDERMAL GROWTH FACTOR RECEPTOR
  Monoclonal Antibodies Against Epidermal Growth Factor Receptor
  Tyrosine Kinase Inhibitors
EPIDERMAL GROWTH FACTOR RECEPTOR–DIRECTED THERAPY IN COMBINATION WITH RADIATION
THERAPIES DIRECTED AGAINST VASCULAR ENDOTHELIAL GROWTH FACTOR
  Antibodies Against Vascular Endothelial Growth Factor
  Other Vascular Endothelial Growth Factor–Directed Therapies
VASCULAR ENDOTHELIAL GROWTH FACTOR–DIRECTED THERAPY IN COMBINATION WITH RADIATION
ADVANCED DISEASE THERAPY TARGETING BOTH VASCULAR ENDOTHELIAL GROWTH FACTOR AND EPIDERMAL GROWTH FACTOR RECPTOR
TARGETING THE STROMA
INSULIN-LIKE GROWTH FACTOR-1–TARGETED THERAPY
POTENTIAL FUTURE THERAPIES
CONCLUSION

Address Correspondence to:

Caio M. Rocha-Lima
Sylvester Comprehensive
Cancer Center
Division of Hematology/
Oncology
Miami, FL
Email: crocha@med.miami.edu

ABSTRACT

Patients with pancreatic cancer normally present with advanced disease that is lethal and notoriously difficult to treat. However, clinical developments over the past decade have been modest and advances are incremental at most. The core drug and the backbone of treatment in all settings of pancreatic adenocarcinoma–adjuvant, locally advanced, and metastatic–remains gemcitabine. The past decade of research focused initially on combining cytotoxic therapies with gemcitabine, and during that time a large number of studies have been published that aimed to target the molecular abnormalities implicated in pancreatic tumor growth, invasion, metastasis, angiogenesis, and resistance to apoptosis. This research is of particular importance, as data suggest that a large number of genetic alterations affect only a few major signaling pathways and processes involved in pancreatic tumorigenesis. Although laboratory results of targeted therapies have been impressive, until now only erlotinib, an epidermal growth factor receptor tyrosine kinase inhibitor, has demonstrated marginal survival benefit in combination with gemcitabine in phase III clinical trials. Even though the failures of targeted therapies in the clinical setting are discouraging, it is reasonable to surmise that major progress will evolve as the molecular biology of pancreatic adenocarcinoma continues to be unraveled. This review describes some of the important developments and targeted agents for pancreatic cancer that have been tested in clinical trials. ***Mt Sinai J Med 77:606–619, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: epidermal growth factor receptor, molecular targeted therapy, pancreatic cancer, tyrosine kinase inhibitors, vascular endothelial growth factor.

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20217

© 2010 Mount Sinai School of Medicine

Adenocarcinoma of the pancreas remains a treatment-refractory malignancy. Although it is only the thirteenth most common cause of cancer worldwide, it is the fifth most common cause of cancer-related death in the United States. The estimated incidence of pancreatic cancer (PC) is 43,140 cases with 36,800 deaths in the United States in 2010.[1] For all the stages combined, the 1- and 5-year survival rates are only 23% and 6%, respectively.[2] The only potential curative treatment offered to patients is a combination of complete tumor removal and adjuvant therapy. Unfortunately, <15% of patients present with resectable disease.[3]

The 5-year survival rate for patients at all stages is 6%. The poor outcome and high recurrence rate after surgery indicate the need for additional treatments either with chemotherapy, radiotherapy, or both. Current adjuvant and neoadjuvant therapies have shown only a small statistically significant improvement in survival over surgery alone.[4,5] A better understanding of the biology of PC has led to development of novel agents targeting pathways of cancer-cell survival. In targeted therapies, tumor-associated or tumor-specific proteins are targeted and their roles in corresponding signal transduction pathways are under investigation.

Many different targeted therapies have been tested in PC.[6] Because gemcitabine has been considered a standard treatment for advanced PC (APC) for the past decade, most clinical trials have explored the combination of gemcitabine plus a biological "targeted" agent. At large this approach has not been successful, and novel strategies are clearly needed. In this article we will review molecular therapies in PC.

KIRSTEN RAT SARCOMA ONCOGENE INHIBITORS

Ninety percent of PCs contain mutations of the Kirsten rat sarcoma (KRAS) proto-oncogene resulting in constitutively activated Ras protein. A critical step in Ras activation is farnesylation of Ras protein. Farnesyl transferase inhibitors (FTIs), compounds that inhibit farnesylation, were studied either as a single agent or in combination with gemcitabine (Table 1). Unfortunately, randomized phase III trials did not show a benefit of adding FTIs to gemcitabine compared with gemcitabine alone (Table 1).

Matrix Metalloproteinases Inhibitors

Matrix metalloproteinases (MMPs) are enzymes that are important for growth, differentiation and repair of normal tissues. Preclinical models demonstrate that aberrant MMP expression contributes to the invasive growth and spread of a variety of solid malignancies.[7] The MMP inhibitors may prevent tumor growth and metastasis by restoring a balance in proteolytic activity. Unfortunately, phase III trials studying the addition of marimastat or BAY 12-9566 to gemcitabine showed no improvement in efficacy results compared with gemcitabine as a single agent (Table 1).

Table 1. *Summary of Phase II/III Trials of MMPI, FTI, mTOR and COX Inhibitors for Advanced or Metastatic Pancreatic Cancer.*

Trial	Phase	Single-Agent Drug Arm	OS (months)	PFS (months)	ORR (%)	SD (%)
Wolpin *et al.*[95] (n = 31)	II	Everolimus	NR	1.9	0	0
MacDonald *et al.*[96] (n = 53)	II	FTI (R115777)	2.6	1.4	NR	NR
Evans *et al.*[97] (n = 113)	II	Marimastat	3.8	NR	8	8
Bramhall *et al.*[98] (n = 414)	III	Marimastat	3.4	1.8	2.8	NR
		Gem	5.5	3.8	25.8	NR
Bramhall *et al.*[99] (n = 239)	III	Marimastat + gem	5.4	3.0	61.0	50.0
		Gem	5.4	3.1	72.0	56.0
Moore *et al.*[100] (n = 277)	III	BAY 12-9566	3.74	1.68	30.9	28.7
		Gem	6.59	3.5	58.2	53.9
			$P < 0.001$	$P < 0.001$		
Cohen *et al.*[101] (n = 20)	II	Tipifarnib	4.8	1.2	0	0
Van Cutsem *et al.*[102] (n = 688)	III	Tipifarnib + gem	6.3	3.7	59.0	53.0
		Gem	6.0	3.6	60.0	52.0
Xiong *et al.*[103] (n = 20)	II	Gem + celecoxib	6.2	3.0	NR	NR
Ferrari *et al.*[104] (n = 42)	II	Gem + celecoxib	9.1	NR	71	62
Dragovich *et al.*[105] (n = 25)	II	Gem + celecoxib	NS	NR	52	35
Lipton *et al.*[106] (n = 21)	II	Gem + celecoxib + irino	18	NR	NR	80

Abbreviations: COX, cyclooxygenase inhibitor; FTI, farnesyl transferase inhibitor; gem, gemcitabine; irino, irinotecan; MMPI, matrix metalloproteinase inhibitors; mTOR, mammalian target of rapamycin; NR, not reported; ORR, overall response rate; OS, median overall survival; PFS, median progression-free survival; SD: stable disease.

Cyclooxygenase-2 Inhibitors

Overexpression of cyclooxygenase-2 (COX-2) correlates with aggressive tumor biology.[8] COX-2 induces angiogenesis and invasiveness, and inhibits apoptosis.[9] Clinical trials combining celecoxib and gemcitabine,[10] or 5-fluorouracil (5-FU),[11] or combining celecoxib with gemcitabine and cisplatin (DDP), were disappointing.[12] No phase III trials have been pursued with COX-2 inhibitors in PC.

PI3K, Akt, and mTOR Inhibitors

A role for the phosphoinositide 3-kinase (PI3K)/Akt/mammalian target of rapamycin (mTOR) pathway in PC biology has been suggested. Pancreatic cancer was found to be associated with activation of the K-RAS protein and downstream activation of RAF/mitogen-activated protein kinase (MAPK), PI3K/Akt/mTOR pathways.[13] Akt2 is amplified and the PI3K-Akt pathway is activated in 20% and 59% of PCs, respectively.[14,15] Inhibitors of PI3K, Akt, and mTOR have demonstrated antitumor activity in PC cells, either alone or in combination.[16,17]

Temsirolimus is an mTOR inhibitor approved for the treatment of renal cell carcinoma, but use of this agent in PC has been limited. The mTOR inhibitor everolimus has been studied in a phase II trial, as a single agen, for second-line therapy in PC patients refractory to gemcitabine.[18] Minimal clinical activity was observed, with no responses and only 21% of the patients having stable disease (SD). Combination therapy is being explored with mTOR inhibitors and other standard or targeted therapies.[19,20] Phase I clinical trials of agents targeting both mTOR and epidermal growth factor receptor (EGFR) are currently enrolling patients.

THERAPIES DIRECTED AGAINST EPIDERMAL GROWTH FACTOR RECEPTOR

A group of transmembrane proteins, EGFRs regulate key processes in the cell, such as proliferation, division, migration, and differentiation. This family has 4 different members: EGFR (HER1 or ErbB1), HER2 (ErbB2), HER3 (ErbB3), and HER4 (ErbB4); all of these share a similar structure.[21] Upon binding to its ligands, homodimerization or heterodimerization with other members of the ErbB family occurs, which leads to phosphorylation of tyrosine residues in its intracellular domain. This process recruits intracellular proteins that cause downstream signaling events through MAPK, PI3K-Akt, and the signal transducer and activator of transcription (STAT) family of proteins[22] (Figure 1).

Mechanisms that lead to inappropriate activation of EGFR include receptor overexpression, activating mutations, overexpression of receptor ligands, and/or loss of their negative regulatory pathways. Overexpression of EGFR and its ligands EGF and TGF-α; are frequently observed in PC.[23,24] High levels of EGFR in PC are associated with early aggressive disease, radiotherapy resistance, decreased survival, and poor prognosis.[21,25,26] Therefore, EGFR inhibition has emerged as a promising therapy for PC. Two different categories of drugs inhibit EGFR activation: the first group contains anti-EGFR monoclonal antibodies (mAbs), which block the activation of the receptor, and the second group, tyrosine kinase inhibitors (TKI), decreases the activity of tyrosine kinase domain by preventing EGFR autophosphorylation.

> *A promising therapy for pancreatic cancer is the inhibition of epidermal growth factor receptor. High levels of epidermal growth factor receptor (EGFR) are associated with early aggressive disease, radiotherapy resistance, decreased survival, and poor prognosis.*

Monoclonal Antibodies Against Epidermal Growth Factor Receptor

Cetuximab

The effectiveness of cetuximab against PC has been tested both in vitro and in vivo. Through EGFR binding, cetuximab induces degradation of the receptor complex, enhances the synthesis of proapoptotic proteins and decreases angiogenesis by decreasing vascular endothelial growth factor (VEGF) and interleukin 8 (IL-8) synthesis.[27]

In a phase II study,[28] 41 patients with EGFR expressing locally advanced, metastatic, or recurrent PC were treated with the combination of gemcitabine and cetuximab. An overall response rate (ORR) of 12.2% was reported, with a further 63.4% of patients achieving disease stabilization. The overall survival (OS) and progression-free survival (PFS) rates after treatment with cetuximab and gemcitabine were felt to justify a follow-up phase III trial, undertaken by the Southwest Oncology Group (SWOG, S0205), comparing the doublet to gemcitabine alone.[29] Unfortunately, no statistically significant differences

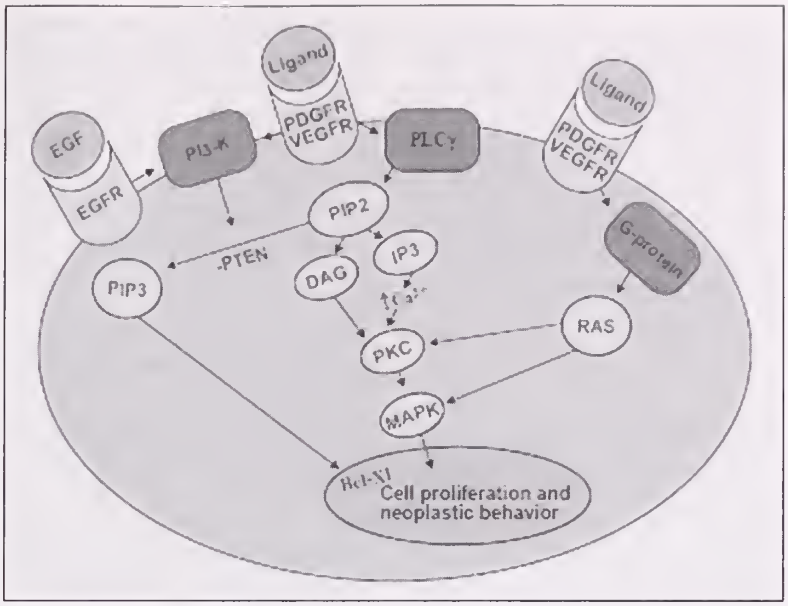


Fig 1. A simplified version of the EGFR and VEGFR activation cascade. Used with permission from *Case Control Journal*. **Abbreviations:** Ca^{2+}, calcium; DAG, diacylglycerol; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; IP3, inositol 1,4,5-trisphosphate; MAPK, mitogen-activated protein kinase; PDGFR, platelet-derived growth factor receptor; PI3-K, phosphatidylinositol 3-kinase; PIP2, phosphatidylinositol 3,4-biphosphate; PIP3, phosphatidylinositol trisphosphate; PLC, phospholipase C; PTEN, phosphatase and tensin homolog; VEGFR, vascular endothelial growth factor receptor.

A phase III trial evaluating the treatment of pancreatic cancer with a combination of gemcitabine and cetuximab versus gemcitabine alone showed no significant difference between the regimens. However, the results of treatment with a combination of cetuximab, gemcitabine, and oxaliplatin have been promising. In a recent phase II trial, the overall response rate of this combination was 24%, with a median progression-free survival and overall survival of 6.9 months and 11.3 months, respectively.

in OS or PFS after 735 patients enrolled was observed. Overall survival was 6 months in the gemcitabine arm and 6.5 months in the gemcitabine + cetuximab arm, for an overall hazard ratio (HR) of 1.09 (95% confidence interval [CI]: 0.93–1.27, $P = 0.14$). Progression-free survival was 3.0 and 3.5 months for the gemcitabine and gemcitabine-cetuximab arms, respectively ($P = 0.058$). The combination of cetuximab, gemcitabine, and oxaliplatin was studied in a phase II trial that enrolled 41 patients with locally APC and metastatic PC (MPC).[30] The ORR was 24%, with median PFS and OS of 6.9 months and 11.3 months, respectively.

Nongemcitabine combinations with cetuximab have also been studied in PC. A randomized phase II trial conducted by the Eastern Cooperative Oncology Group (ECOG) included 87 eligible patients with MPC to docetaxel and irinotecan with or without cetuximab.[31–33] Unexpectedly, the patients treated with cetuximab had a shorter median survival duration of 5.3 months compared with 6.5 months.[32] The results of a single-arm, phase II study assessing the efficacy of cetuximab combined with ixabepilone was recently reported.[34] Of 54 patients enrolled, 93% had MPC. Six- and 12-month survival rates were 62% and 40%, respectively. The median PFS was 3.9 months (95% CI: 2.6–4.4 months). These results were similar to a previous phase II trial with single-agent ixabepilone in MPC.[35] No further studies with ixabepilone and cetuximab in PC have been planned.

Matuzumab

Matuzumab has similar anti-EGFR affinity compared with cetuximab. Safety after combined therapy with matuzumab + gemcitabine was evaluated in a phase I trial that included patients with APC. Acceptable adverse events were observed when matuzumab and gemcitabine are used together. Although it was not the primary endpoint, according to the investigators, partial response (PR) or SD was observed in 8 of 12 patients included in this trial.[36]

Trastuzumab

In 10%–82% of PCs, ErbB2/HER2 has been found to be overexpressed. Kimura *et al.*[37] tested cytotoxicity of HER2-mAb trastuzumab in 4 PC cell lines and reported no inhibitory, synergistic, or additive effect when combined with gemcitabine. However, antibody-dependent cell-mediated cytotoxicity (ADCC) was observed in 3 of 4 cell lines and was proportional to HER2 expression. Dual inhibition of EGFR and HER2 has also been reported by Larbouret *et al.*,[38] who showed that combination of matuzumab and trastuzumab significantly decreased xenograft tumor sizes or induced more complete remissions (CR) as compared with either antibody alone, and prolonged survival in mice inoculated with PC cells. The same author recently published the results of an in vitro study comparing the tumor growth inhibitory effect of combined cetuximab and trastuzumab versus gemcitabine on nude mice bearing human pancreatic carcinoma xenografts. This preclinical study demonstrated a significant improvement of survival and tumor regression in mice treated with the 2 mAbs compared with gemcitabine, independently of the K-Ras status.[39] Safran *et al.*[40] conducted a study assessing the ORR of trastuzumab combined with gemcitabine in patients with MPC that overexpress HER-2/neu (only 16% of the patients tested overexpressed HER-2 neu). Confirmed PR was observed in 6% of 34 enrolled patients; the median survival was 7 months, and the 1-year survival was 19%. Combination treatment with trastuzumab and other tumor inhibitory agents, such as the histone deacetylase inhibitor vorinostat, the farnesyltransferase inhibitor tipifarnib, or IL-12, is being tested in clinical trials.[41–43]

> *Safran* et al., *studying the treatment of patients with metastatic pancreatic cancer that overexpresses the EGFR variant HER-2/neu, evaluated the effect of trastuzumab combined with gemcitabine. A confirmed partial response was observed in 6% of 34 enrolled patients; the median survival was 7 months, and the 1-year survival was 19%.*

Tyrosine Kinase Inhibitors

Erlotinib

Erlotinib is the only TKI approved by the US Food and Drug Administration for patients with APC who have not received previous chemotherapy. Preclinical synergy with gemcitabine and erlotinib in inducing apoptosis in pancreatic xenograft models was demonstrated,[44] and a phase I study established the dose of erlotinib for single-agent daily dosing to be 150 mg/day.[45] In a phase III trial performed by the National Cancer Institute of Canada Clinical Trials Group (NCIC CTG),[46] 569 patients with APC or MPC were randomly assigned to receive a lower dose of erlotinib 100 mg/day + gemcitabine versus gemcitabine alone. The study showed statistically significant improvements in OS (6.37 months in the erlotinib arm and 5.91 months in the control arm, $P = 0.038$). As in studies of anti-EGFR agents in colorectal cancer, presence of rash was associated with a higher likelihood of achieving disease control ($P = 0.05$). An additional small phase II trial combining erlotinib with capecitabine has shown a 10% ORR. The median survival duration was

> *In a phase III trial performed by the National Cancer Institute of Canada Clinical Trials Group, 569 patients with advanced pancreatic cancer or metastatic pancreatic cancer were randomly assigned to receive a relatively low dose of erlotinib 100 mg/day + gemcitabine versus gemcitabine alone. The study showed statistically significant improvements in overall survival (6.37 months in the erlotinib arm and 5.91 months in the control arm, $P = 0.038$).*

6.5 months,[47] similar to results of other published phase II trials of second-line therapy.[17,18]

The efficacy of EGFR inhibitors, including erlotinib, may be impacted by KRAS mutation status.[49] Boeck *et al.*[50] examined the tissue tumor samples of 204 patients treated with gemcitabine + erlotinib followed by capecitabine, versus capecitabine + erlotinib followed by gemcitabine in APC. Of the tumors, 123 (70%) harbored a somatic KRAS mutation (mut). Patients with KRAS wild type (wt, $n = 53$) had an improved OS (wt: 8.0 months versus mut: 6.6 months, $P = 0.011$). The impact of KRAS mutations on the results of combination regimens could be due to the intrinsic prognostic impact of KRAS mutations, independent of the treatment applied.

> *The efficacy of EGFR inhibitors, including erlotinib, may be impacted by Kirsten rat sarcoma (KRAS) viral oncogene mutation status. Boeck* et al. *examined the tissue tumor samples of 204 advanced pancreatic cancer patients treated with gemcitabine + erlotinib followed by capecitabine, versus capecitabine + erlotinib followed by gemcitabine. Of the tumors, 123 (70%) harbored a somatic Kirsten rat sarcoma (KRAS) proto-oncogene mutation. Patients with KRAS wild type ($n = 53$) had an improved overall survival (wild type: 8.0 months versus mutation: 6.6 months, $P = 0.011$).*

Gefitinib

Gefitinib has an inhibitory effect in cell growth in vitro and in vivo that is dose dependent. As observed with other EGFR inhibitors, gefitinib is a good chemosensitizer and radiosensitizer because it suppresses the DNA repair induced after these treatments.[51] The combination of gefitinib and gemcitabine has also been studied in a single-arm phase II study in Europe. Fifty-three patients with APC or MPC were treated with gefitinib and gemcitabine. The 6-month PFS was 30%, with a median PFS of 4.1 months. The 1-year survival rate was 27%, with a median survival of 7.3 months.[52] These results are comparable with those from the erlotinib and gemcitabine arm in the NCIC CTG trial,[46] but have not been confirmed in a randomized trial.

EPIDERMAL GROWTH FACTOR RECEPTOR–DIRECTED THERAPY IN COMBINATION WITH RADIATION

It has been previously demonstrated that EGFR activation contributes to radiation resistance,[53,54] and in preclinical models this enhanced tumor-cell survival and radiation resistance may be overcome by inhibition of EGFR.[55–57] Therefore, the role of EGFR-directed therapy in combination with radiation therapy has been under active investigation. A recent phase I study of erlotinib in combination with twice-weekly gemcitabine and radiotherapy for APC showed this combination to be well-tolerated, with a PR rate of 35% and SD rate of 53% in 17 assessable patients.[58] Final results of a phase II trial[59] examining the addition of cetuximab to gemcitabine and intensity-modulated radiation therapy in patients with APC were recently presented. Sixty-eight chemotherapy-naïve patients received intensity-modulated radiation therapy with weekly gemcitabine (300 mg/m^2) and cetuximab (400 mg/m^2 loading dose, then 250 mg/m^2). Patients were randomized following completion of radiotherapy to weekly gemcitabine alone (1000 mg/m^2) for 4 weeks, or weekly gemcitabine (1000 mg/m^2) and cetuximab (250 mg/m^2) over 12 weeks. All patients were subsequently treated with weekly gemcitabine alone. One patient had a CR, 23 had PR, 41 had SD, and only 3 had progressive disease (PD). Fourteen patients went on to have complete surgical resections. Although these results appear promising, a randomized trial comparing cetuximab, gemcitabine, and radiation with gemcitabine and radiation has yet to be performed. A phase III study of the Radiotherapy Oncology Group/SWOG/NCIC/European Organisation for Research and Treatment of Cancer involving 2 randomizations opened in summer 2009 to determine whether the addition of erlotinib to gemcitabine adjuvant chemotherapy improves survival as compared with gemcitabine alone following R0 or R1 resection of head of PC. Second randomization is being undertaken to determine whether the use of concurrent fluoropyrimidine and radiotherapy following adjuvant gemcitabine-based chemotherapy further enhances survival for such patients who are without evidence of PD after 5 cycles of gemcitabine-based chemotherapy. Tables 2, 3, and 4 summarize the

Table 2. *Summary of Phase II/III Trials of EGFR-Targeted Therapies for Advanced or Metastatic Pancreatic Cancer.*

Trial	Phase	Combination	OS (months)	PFS (months)	ORR (%)	SD (%)
Xiong *et al.*[28] (n = 61)	II	Gem + cetuximab	7.1	3.8	12.2	63.4
Rocha-Lima *et al.*[34] (n = 54)	II	Cetuximab + ixabepilone	7.6	3.9	13.0	77.0
Fountzilas *et al.*[52] (n = 53)	II	Gem + gefitinib	7.3	4.1	NR	NR
Burtness *et al.*[32] (n = 91)	II	Irino + doc	6.5	3.9	NR	NR
		Irino + doc + cetuximab	5.3	4.5		
Moore *et al.*[46] (n = 569)	III	Gem + erlotinib	6.37	3.75	57	48.9
		Gem	5.91	3.55	49.2	41.2
			$P = 0.038$	$P = 0.004$		
Philip *et al.*[29] (n = 735)	III	Gem + cetuximab	6.5	3.5	12.0	NR
		Gem	6.0	3.0	14.0	
			$P = 0.14$	$P = 0.058$		

Abbreviations: Doc, docetaxel; EGFR, epidermal growth factor receptor; gem, gemcitabine; irino, irinotecan; NR, not reported; ORR, overall response rate; OS, median overall survival; PFS, median progression-free survival; SD: stable disease.

Table 3. *Summary of Phase II/III Trials of VEGFR-Targeted Therapies for Advanced or Metastatic Pancreatic Cancer.*

Trial	Phase	Combination	OS (months)	PFS (months)	ORR (%)	SD (%)
Fogelman *et al.*[65] (n = 50)	II	GemOx + bev	12.1	NR	NR	39
Kim *et al.*[67] (n = 82)	II	GemOx + bev	8.1	5.7	11.3	NR
Kindler *et al.*[64] (n = 602)	III	Gem + bev	5.7	4.8	53.8	40.7
		Gem	6.0	4.3	47.0	35.7
			$P =$ NS	$P =$ NS		

Abbreviations: Bev, bevacizumab; gem, gemcitabine; GemOx, gemcitabine 1000 mg/m^2/100 minutes/day 1 + oxaliplatin 100 mg/m^2/day 2 every 14 days; NR, not reported; NS, not significant; ORR, overall response rate; OS, median overall survival; PFS, median progression-free survival; SD, stable disease; VEGFR, vascular endothelial growth factor receptor.

Table 4. *Summary of Phase II/III Clinical Trials of Combined EGFR- and VEGFR-Targeted Therapies for Advanced or Metastatic Pancreatic Cancer.*

Trial	Phase	Combination	OS (months)	PFS (months)	ORR (%)	SD (%)
Ko *et al.*[107] (n = 57)	II	Gem + cetuximab + bev	NR	3.5	10.7	29
		Cetuximab + bev		1.8	0	24
Kindler *et al.*[108] (n = 139)	II	Gem + bev + erlotinib	7.8	5.0	23	49
		Gem + bev + cetuximab	7.2	5.1	18	45
Vervenne *et al.*[75] (n = 607)	III	Gem + erlotinib	6.0	3.6	8.6	NR
		Gem + erlotinib + bev	7.1	4.6	13.5	

Abbreviations: Bev, bevacizumab; EGFR, epidermal growth factor receptor; gem, gemcitabine; NR, not reported; ORR, overall response rate; OS, median overall survival; PFS, median progression-free survival; SD, stable disease; VEGFR, vascular endothelial growth factor receptor.)

phase II/III trials of EGFR- and VEGFR-targeted therapies for APC or MPC.

THERAPIES DIRECTED AGAINST VASCULAR ENDOTHELIAL GROWTH FACTOR

Angiogenesis is essential for solid-tumor growth, and is principally mediated by the VEGF family of proteins and receptors.[60] The VEGF receptor (VEGFR) is similar to EGFR because after it binds to its ligand, the VEGFR forms a dimer, and the tyrosine kinase becomes autophosphorylated and activates the intracellular pathways, such as MAPK and PI3K.[61] A study also suggested that VEGF may function as an autocrine growth factor

> *Vascular endothelial growth factor is overexpressed in >90% of pancreatic cancers and is, therefore, an appealing target for therapy.*

for PC cells.[62] Vascular endothelial growth factor is overexpressed in >90% of PCs and is, therefore, an appealing target for therapy.

Antibodies Against Vascular Endothelial Growth Factor

Bevacizumab

Bevacizumab is the only humanized antibody against VEGF that has been used in clinical trials. Gemcitabine and bevacizumab were combined in a phase II study with 52 patients having stage IV PC; 19% of the patients had a PR, whereas 48% had SD. Median OS was 8.8 months, with median time to progression of 5.4 months and a 6-month survival rate of 77%.[63] However, the results of the CALBG phase III randomized trial of gemcitabine with or without bevacizumab did not reveal any improvement in survival upon addition of bevacizumab.[64] Median OS and PFS in bevacizumab versus control arm were 5.7 months and 6.0 months, respectively (95% CI: 4.9–6.5 months versus 5.0–6.9 months) and 4.8 and 4.3 months, respectively (95% CI: 4.3–5.7 months versus 3.8–5.6 months).

> *Though a phase II study showed promising results in the treatment of stage IV pancreatic cancer with a combination of gemcitabine and bevacizumab, an anti-VEGF antibody, a phase III randomized trial of gemcitabine with or without bevacizumab failed to show that the latter produced any improvement in survival.*

Two phase II trials were presented at the American Society of Clinical Oncology's 2009 Gastrointestinal Cancers Symposium evaluating the efficacy of gemcitabine and oxaliplatin (GemOx) + bevacizumab in 50 patients with APC. This triple-drug combination has achieved 1- and 2-year survival rates of 40% and 16%, respectively, with a high ORR of 39%.[65] Also, this regimen demonstrated a higher ORR and longer median survival compared with previously reported gemcitabine and oxaliplatin study.[66] Of note, there was a correlation between CA 19-9 levels and median survival. The second phase II trial[67] assessing the same triple combination included 82 patients with APC. This study showed 6-month survival of 65% (95% CI: 53.5%–75.3%), median survival of 8.1 months (95% CI: 6.5–9.3 months), and median TTP of 5.7 months (95% CI: 4.4–6.4 months). Moreover, Javle *et al.*[68] reported the clinical efficacy of gemcitabine, capecitabine, and bevacizumab combination in 50 patients with APC. The treatment consisted of bevacizumab 15 mg/kg, capecitabine 1300 mg/m^2 daily for 2 weeks, and gemcitabine 1000 mg/m^2 weekly 2 times, with the cycle to be repeated every 21 days. Radiological response rate was 22%; PFS and OS were 5.8 and 9.8 months, respectively.

Other Vascular Endothelial Growth Factor–Directed Therapies

The failure of bevacizumab in therapeutic trials for PC highlighted the need for angiogenic inhibitors that could target the VEGF pathway differently than bevacizumab and/or have better access to the tumor environment than an antibody.

Sorafenib

Sorafenib is a multitargeted kinase inhibitor (Raf-kinase, VEGF, and platelet-derived growth factor receptor). It was approved in 2005 for the treatment of advanced renal cell carcinoma. One phase I study showed good tolerance after treatment with gemcitabine and sorafenib in PC: 56% of the patients had SD, and no PR was seen.[69] However, a phase II study concluded that, although well-tolerated, it was inactive in patients with APC.[70]

Axitinib

Axitinib at high concentrations is an orally active inhibitor of both VEGFR and related tyrosine kinase receptors. A median survival of 6.9 months was reported for axitinib combined with gemcitabine compared with 5.6 months for gemcitabine alone in a phase II trial in patients with APC, but this finding was not significant.[71] Phase III testing has been conducted in APC; however, a press release[72] announced the discontinuation of the study based on an interim analysis that found no evidence of improvement in the primary endpoint of survival in patients treated with axitinib and gemcitabine, compared with gemcitabine alone.

Aflibercept

Aflibercept, a recombinant fusion protein that functions as a VEGF inhibitor, is another novel agent being tested as a global study in a phase III trial of patients treated with gemcitabine for MPC with performance status of 0–1, and was deemed futile in a press release in November 2009.

VASCULAR ENDOTHELIAL GROWTH FACTOR–DIRECTED THERAPY IN COMBINATION WITH RADIATION

Bevacizumab and radiation can be combined with acceptable toxicity. Therapy with bevacizumab, capecitabine, and radiotherapy was safe in a phase I study and did not produce a significant increase in toxicity.[73] The RTOG has recently published the results of the phase II trial evaluating this regimen.[74] Patients with APC without duodenal invasion were treated with 50.4 Gy radiation to gross tumor with concurrent capecitabine 825 mg/m^2 orally twice daily on days of radiation and bevacizumab 5 mg/kg on days 1, 15, and 29, followed by maintenance gemcitabine 1 g/m^2 weekly for 3 weeks and bevacizumab 5 mg/kg every 2 weeks, both in 4-week cycles until progression. The median and 1-year survival rates were 11.9 months (95% CI: 9.9–14.0 months) and 47% (95% CI: 36%–57%), respectively. Median PFS was 8.6 months (95% CI: 6.9–10.5 months) and ORR was 26%. Overall, 35.4% of patients had treatment-related gastrointestinal toxicity ≥ grade III. This study resulted in a similar median survival to previous RTOG studies in patients with APC.

ADVANCED DISEASE THERAPY TARGETING BOTH VASCULAR ENDOTHELIAL GROWTH FACTOR AND EPIDERMAL GROWTH FACTOR RECPTOR

Following up on the NCIC CTG trial of erlotinib and gemcitabine, the Avastin/Tarceva (AVITA) study was conducted, in which[75] 607 chemotherapy-naïve patients with MPC were randomized to receive placebo or bevacizumab (5 mg/kg every 2 weeks) in combination with gemcitabine (1000 mg/m^2 weekly for 7 weeks, followed by a 1-week rest, then 1000 mg/m^2 weekly for 3 weeks, every 4 weeks) and erlotinib (100 mg daily). At the final analysis, 454 patients had died (233 in the placebo arm; 221 in the bevacizumab arm). The median OS was not significantly different between the 2 groups (6.0 months for the placebo arm versus 7.1 months in the bevacizumab arm (HR: 0.89, 95% CI: 0.74–1.07); however, the median PFS was significantly longer for the bevacizumab-treated patients (3.6 months for the placebo arm versus 4.6 months for the bevacizumab arm (HR: 0.73, 95% CI: 0.61–0.86, $P < 0.01$).

TARGETING THE STROMA

Recent studies suggest that the stroma blocks drug penetration contributing to tumor resistance. Pancreatic cancer cells and surrounding stroma overexpress secreted protein acidic and rich in cysteine (SPARC). Some researchers consider the stromal depletion hypothesis already clinically validated, at least for albumin-bound paclitaxel (nab-paclitaxel, NP). Nab-paclitaxel increases tumor accumulation of paclitaxel through binding of albumin to SPARC. Nab-paclitaxel + gemcitabine demonstrated clinical benefit in phase I/II trials.[76] A total of 67 stage IV PC patients were enrolled. Of 58 evaluable patients, 23 achieved a PR and 22 SD. For all patients, the median time for PFS was 6.9 months, and OS was 10.3 months. Abraxis BioScience launched a phase III clinical trial in 2009 in MPC comparing NP and gemcitabine versus gemcitabine alone, and is currently enrolling patients. Single-agent NP has shown activity in PC. Hosein *et al.* demonstrated a clinical benefit of NP in a phase II trial in patients with APC who have progressed on gemcitabine-based therapy.[77] The 6-month and median OS were 63% (95% CI: 38%–80%) and 7.3 months (95% CI: 2.8–not attained), respectively. The median PFS was 1.7 months (95% CI: 1.5–3.4 months). One patient had a confirmed PR, and 6 (32%) had SD. The remaining 12 patients (63%) had PD on or before the first response assessment. Whether SPARC expression is a predictive biomarker needs further investigation.

Hosein et al. *demonstrated a clinical benefit of albumin-bound paclitaxel in a phase II trial in patients with APC who have progressed on gemcitabine-based therapy. The 6-month and the median OS were 63% (95% CI: 38%–80%) and 7.3 months (95% CI: 2.8– not attained), respectively.*

Another promising stromal targets is hedgehog signaling. Three mammalian hedgehog homolog (HH) proteins have been identified–desert hedgehog homolog (DHH), Indian hedgehog homolog (IHH), and sonic hedgehog homolog (SHH)–activation of which is controlled by 2 transmembrane proteins, the tumor-suppressor PTC1 protein and the oncogenic sMo protein. In 70% of human PC, SHH is expressed; IHH expression is increased 35-fold in PC cells compared with in normal tissues.[78] Preclinical studies have demonstrated that cyclopamine (sMo inhibitor) can enhance sensitivity to radiotherapy and chemotherapy and suppress metastatic spread[79,80] as well as improve antitumor activity when combined with an EGFR inhibitor.[81] Additionally, Ogino *et al.*[82] evaluated hedgehog signaling pathway as a therapeutic target for dendritic cell–based immunotherapy against patients with solid tumors and demonstrated that cyclopamine could suppress increased VEGF secretion of Mo-DCs in an hypoxic environment. Tuveson *et al.* from the United Kingdom showed that gemcitabine concentration in the tumors rose by 60% if preceded by a HH inhibitor, and the combination more than doubled mouse survival compared with gemcitabine alone. Two clinical trials of the Genentech HH inhibitor GDC-0449 in PC are under way, and a third was pending as of early 2010.

INSULIN-LIKE GROWTH FACTOR-1–TARGETED THERAPY

The insulin-like growth factor 1 receptor (IGFR-1), a transmembrane receptor tyrosine kinase, is overexpressed in 64% of PCs.[83] Human IGFR-1 have been reported to increase the antitumor effects of gemcitabine. Kindler *et al.*[84] designed a 3-arm, placebo-controlled, randomized phase II study to estimate the efficacy of adding AMG 479 (IGFR-1; A) or conatumumab (death receptor 5 agonist; C) to gemcitabine (G) in untreated MPC patients. Patients were randomly allocated 1:1:1 to Arm 1, C + G (double-blinded); Arm 2, A + G (open label due to anticipated thrombocytopenia and hyperglycemia); or Arm 3, P + G (double-blinded). The addition of C or A to G was well-tolerated and resulted in trends toward longer PFS, improved 6-month OS, and higher rates of SD. Also, Javle *et al.*[85] reported the results of an ongoing phase I/II study of MK-0646, the IGFR-1 in combination with gemcitabine (G) (arm A), or gemcitabine + erlotinib (E) plus MK-0646 (arm B), or G + E (arm C) for advanced PC. Of 28 enrolled patients, 6 had PR (including 1 hepatic CR), 10 PD, 8 SD, and 5 sustained PRs (range, 20+–44+ weeks). Median time to progression was 8 weeks (range, 4–44+weeks). The SWOG has completed enrolling patients into a phase I/II trial of gemcitabine and erlotinib therapy with or without the addition of an IGFR-1, IMC-A12,. in chemo-naïve patients with MPC. Study completion in 2011 is anticipated.

POTENTIAL FUTURE THERAPIES

Some epithelial tumors, including PC, have intrinsic mutations that make them primarily resistant or eventually resistant during the course of the treatment. The most well-defined mechanisms of resistance involve activation of downstream mediators, PI3K-Akt; loss of PTEN, a lipid phosphatase and tumor suppressor protein; amplification of MET, a protooncogene that encodes a transmembrane tyrosine kinase receptor; KRAS mutations; or by conformational alterations in the adenosine triphosphate binding domain of the receptors. Therefore, future areas of development of new therapeutics include other kinase inhibitors,[86] PI3K-Akt signaling inhibition, Poly ADP ribose polymerase (PARP) inhibition,[87] c-MET inhibition, and antagonists of the tumor microenvironment, to name a few.[88]

There have also been many studies characterizing the cause of drug resistance in PC with respect to discovering therapeutic mechanisms to specifically reach the tumorigenic cancer stem cells themselves.[89,90] Evidence for existence of CD133+ PC stem cells was first reported by Hermann *et al.*[91] The authors found that the subpopulation of CD133+ cells were exclusively tumorigenic and highly resistant to gemcitabine. A specific ligand of the CXCR4 receptor, SDF-1, was found to be the strongest inducer of migration for CD133+ cancer cells in vitro.[92] Recently, a report demonstrated that when highly tumorigenic cancer cells undergo epithelial to mesenchymal transition (EMT), they begin to express tumor-initiator markers.[93] Pino *et al.*[94] showed that EMT and EMT-related biomarkers were specifically associated with EGFR dependency and sensitivity to erlotinib in PC tumor cell lines. A better understanding of the biological behavior of these cells may be useful for selecting patients at the highest likelihood of benefit from EGFR TKIs and improve outcomes in patients with this devastating disease.

CONCLUSION

Pancreatic cancer cells are resistant to conventional therapies due to overexpression of different mutations that prevent cell death. The recent introduction of targeted therapies against PC raised hopes in the treatment of patients with advanced or metastatic disease. Unfortunately, the target agents studied to date have fallen short of the expectations due to emerging resistance from multiple mutations that bypass one pathway through another pathway. The preclinical evidence using combined targeted therapies or blocking the activation of different pathways through common downstream effectors is promising. Because the current evidence is limited to small nonrandomized trials or phase I studies that evaluate the safety of the compounds, clinical utility remains to be established. Additional studies trying to identify predictors of response and to compare these new therapies with the current standard of care are needed.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. National Cancer Institute. Common cancers and 2010 statistics. http://www.cancer.gov/cancertopics/commoncancers. Published 2010. Accessed on July 14, 2010.
2. American Cancer Society. Cancer Facts and Figures 2005 http://www.cancer.org/Research/CancerFacts Figures/cancer-facts-figures-2005
3. Muller MW, Friess H, Köninger J, et al. Factors influencing survival after bypass procedures in patients with advanced pancreatic adenocarcinomas. *Am J Surg* 2008; 195: 221–228.
4. Postier RG. The challenge of pancreatic cancer. *Am J Surg* 2003; 186: 579–582.
5. Bilimoria KY, Bentrem DJ, Tomlinson JS, et al. Quality of pancreatic cancer care at Veterans Administration compared with non-Veterans Administration hospitals. *Am J Surg* 2007; 194: 588–593.
6. US National Institutes of Health. ClinicalTrials.gov. http://clinicaltrials.gov/. Accessed on July 14, 2010.
7. Chambers AF, Matrisian LM. Changing views of the role of matrix metalloproteinases in metastasis. *J Natl Cancer Inst* 1997; 89: 1260–1270.
8. Wang HX, Chen QK. Expression and significance of cyclooxygenase-2 in human pancreatic carcinomas [in Chinese]. *Ai Zheng* 2003; 22: 649–652.
9. Zhou XC, Tang CW, Liu CL, et al. Effects of rofecoxib on angiogenesis of pancreatic cancer xenograft in nude mice [in Chinese]. *Ai Zheng* 2004; 23: 376–380.
10. Xiong HQ, Hess KR, Kayaleh OR, et al. A phase II trial of gemcitabine and celecoxib for metastatic pancreatic cancer. *J Clin Oncol* 2005; 23:(suppl 16S):abstract 4174.
11. Milella M, Gelibter A, Di Cosimo S, et al. Pilot study of celecoxib and infusional 5-fluorouracil as second-line treatment for advanced pancreatic carcinoma. *Cancer* 2004; 101: 133–138.
12. El-Rayes BF, Zalupski MM, Shields AF, et al. Phase II study of celecoxib, gemcitabine, and cisplatin in advanced pancreatic cancer. *Invest New Drugs* 2005; 23: 583–590.
13. Hezel AF, Kimmelman AC, Stanger BZ, et al. Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev* 2006; 20: 1218–1249.
14. Ruggeri BA, Huang L, Wood M, et al. Amplification and overexpression of the *AKT2* oncogene in a subset of human pancreatic ductal adenocarcinomas. *Mol Carcinog* 1998; 21: 81–86.
15. Schlieman MG, Fahy BN, Ramsamooj R, et al. Incidence, mechanism and prognostic value of activated Akt in pancreas cancer. *Br J Cancer* 2003; 89: 2110–2115.
16. Bondar VM, Sweeney-Gotsch B, Andreeff M, et al. Inhibition of the phosphatidylinositol 3′-kinase-AKT pathway induces apoptosis in pancreatic carcinoma cells in vitro and in vivo. *Mol Cancer Ther* 2002; 1: 989–997.
17. Bruns CJ, Koehl GE, Guba M, et al. Rapamycin-induced endothelial cell death and tumor vessel thrombosis potentiate cytotoxic therapy against pancreatic cancer. *Clin Cancer Res* 2004; 10: 2109–2119.
18. Ito D, Fujimoto K, Mori T, et al. In vivo antitumor effect of the mTOR inhibitor CCI-779 and gemcitabine in xenograft models of human pancreatic cancer. *Int J Cancer* 2006; 118: 2337–2343.
19. Azzariti A, Porcelli L, Gatti G, et al. Synergic antiproliferative and antiangiogenic effects of EGFR and mTOR inhibitors on pancreatic cancer cells. *Biochem Pharmacol* 2008; 75: 1035–1044.
20. Tuncyurek P, Mayer JM, Klug F, et al. Everolimus and mycophenolate mofetil sensitize human pancreatic cancer cells to gemcitabine *in vitro:* a novel adjunct to standard chemotherapy? *Eur Surg Res* 2007; 39: 380–387.
21. Zhang Y, Banerjee S, Wang Z, et al. Antitumor activity of epidermal growth factor receptor-related protein is mediated by inactivation of ErbB receptors and nuclear factor-kappaB in pancreatic cancer. *Cancer Res* 2006; 66: 1025–1032.
22. Wong HH, Lemoine NR. Pancreatic cancer: molecular pathogenesis and new therapeutic targets. *Nat Rev Gastroenterol Hepatol* 2009; 6: 412–422.
23. Korc M, Chandrasekar B, Yamanaka Y, et al. Overexpression of the epidermal growth factor receptor in human pancreatic cancer is associated with concomitant increases in the levels of epidermal growth factor and transforming growth factor alpha. *J Clin Invest* 1992; 90: 1352–1360.
24. Bloomston M, Bhardwaj A, Ellison EC, et al. Epidermal growth factor receptor expression in pancreatic carcinoma using tissue microarray technique. *Dig Surg* 2006; 23: 74–79.
25. Pino MS, Shrader M, Baker CH, et al. Transforming growth factor alpha expression drives constitutive epidermal growth factor receptor pathway activation and sensitivity to gefitinib (Iressa) in human pancreatic cancer cell lines. *Cancer Res* 2006; 66: 3802–3812.
26. Bowers G, Reardon D, Hewitt T, et al. The relative role of ErbB1-4 receptor tyrosine kinases in radiation signal transduction responses of human carcinoma cells. *Oncogene* 2001; 20: 1388–1397.
27. Overholser JP, Prewett MC, Hooper AT, et al. Epidermal growth factor receptor blockade by antibody IMC-C225 inhibits growth of a human pancreatic carcinoma xenograft in nude mice. *Cancer* 2000; 89: 74–82.
28. Xiong HQ, Rosenberg A, LoBuglio A, et al. Cetuximab, a monoclonal antibody targeting the epidermal growth factor receptor, in combination with gemcitabine for advanced pancreatic cancer:

a multicenter phase II trial. *J Clin Oncol* 2004; 22: 2610–2616.

29. Philip PA, Benedetti J, Corless JL, et al. Phase III study of gemcitabine plus cetuximab versus gemcitabine in patients with locally advanced or metastatic pancreatic adenocarcinoma: SWOG S0205 study. *J Clin Oncol* 2010; 28: 3605–3610.
30. Merchan JR, Venkatraman A, Macintyre J, et al. Phase II study of gemcitabine, oxaliplatin and cetuximab in advanced pancreatic cancer. Presented at: ASCO Gastrointestinal Cancers Symposium 2009.
31. Bleickardt E, Argiris A, Rich R, et al. Phase I dose escalation trial of weekly docetaxel plus irinotecan in patients with advanced cancer. *Cancer Biol Ther* 2002; 1: 646–651.
32. Burtness BA, Berlin JD, Liles DK, et al. Phase II ECOG trial of irinotecan-docetaxel with or without cetuximab in metastatic pancreatic cancer: updated survival and CA19-9 results. *J Clin Oncol* 2008; 26:(suppl):abstract 4642.
33. Burtness B, Thomas L, Sipples R, et al. Phase II trial of weekly docetaxel/irinotecan combination in advanced pancreatic cancer. *Cancer J* 2007; 13: 257–262.
34. Rocha-Lima CM, Lin EH, Kim GP, et al. Phase II trial of ixabepilone plus cetuximab as first-line therapy for advanced pancreatic carcinoma. *J Clin Oncol* 2010; 28:(suppl 15S):abstract 4086.
35. Whitehead RP, McCoy S, Rivkin SE, et al. A phase II trial of epothilone B analogue BMS-247550 (NSC #710428) ixabepilone, in patients with advanced pancreas cancer: a Southwest Oncology Group study. *Invest New Drugs* 2006; 24: 515–520.
36. Graeven U, Kremer B, Sudhoff T, et al. Phase I study of the humanised anti-EGFR monoclonal antibody matuzumab (EMD 72000) combined with gemcitabine in advanced pancreatic cancer. *Br J Cancer* 2006; 94: 1293–1299.
37. Kimura K, Sawada T, Komatsu M, et al. Antitumor effect of trastuzumab for pancreatic cancer with high HER-2 expression and enhancement of effect by combined therapy with gemcitabine. *Clin Cancer Res* 2006; 12: 4925–4932.
38. Larbouret C, Robert B, Navarro-Teulon I, et al. In vivo therapeutic synergism of anti-epidermal growth factor receptor and anti-HER2 monoclonal antibodies against pancreatic carcinomas. *Clin Cancer Res* 2007; 13: 3356–3362.
39. Larbouret C, Robert B, Bascoul-Mollevi C, et al. Combined cetuximab and trastuzumab are superior to gemcitabine in the treatment of human pancreatic carcinoma xenografts. *Ann Oncol* 2010; 21: 98–103.
40. Safran H, Iannitti D, Ramanathan R, et al. Herceptin and gemcitabine for metastatic pancreatic cancers that overexpress HER-2/neu. *Cancer Invest* 2004; 22: 706–712.
41. Cengel KA, Voong KR, Chandrasekaran S, et al. Oncogenic K-Ras signals through epidermal growth factor receptor and wild-type H-Ras to promote radiation survival in pancreatic and colorectal carcinoma cells. *Neoplasia* 2007; 9: 341–348.
42. Kumagai T, Wakimoto N, Yin D, et al. Histone deacetylase inhibitor, suberoylanilide hydroxamic acid (Vorinostat, SAHA) profoundly inhibits the growth of human pancreatic cancer cells. *Int J Cancer* 2007; 121: 656–665.
43. Parihar R, Nadella P, Lewis A, et al. A phase I study of interleukin 12 with trastuzumab in patients with human epidermal growth factor receptor-2-overexpressing malignancies: analysis of sustained interferon gamma production in a subset of patients. *Clin Cancer Res* 2004; 10: 5027–5037.
44. Ng SS, Tsao MS, Nicklee T, et al. Effects of the epidermal growth factor receptor inhibitor OSI-774, Tarceva, on downstream signaling pathways and apoptosis in human pancreatic adenocarcinoma. *Mol Cancer Ther* 2002; 1: 777–783.
45. Hidalgo M, Siu LL, Nemunaitis J, et al. Phase I and pharmacologic study of OSI-774, an epidermal growth factor receptor tyrosine kinase inhibitor, in patients with advanced solid malignancies. *J Clin Oncol* 2001; 19: 3267–3279.
46. Moore MJ, Goldstein D, Hamm J, et al. Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* 2007; 25: 1960–1966.
47. Kulke MH, Blaszkowsky LS, Ryan DP, et al. Capecitabine plus erlotinib in gemcitabine-refractory advanced pancreatic cancer. *J Clin Oncol* 2007; 25: 4787–4792.
48. Oh L, Lee K, Sohn C, et al. A phase II trial of erlotinib in combination with gemcitabine and capecitabine in previously untreated metastatic/recurrent pancreatic cancer: combined analysis with translational research. *J Clin Oncol* 2009; 26:(suppl):abstract 4638.
49. Laghi L, Orbetegli O, Bianchi P, et al. Common occurrence of multiple K-RAS mutations in pancreatic cancers with associated precursor lesions and in biliary cancers. *Oncogene* 2002; 21: 4301–4306.
50. Boeck SH, Vehling-Kaiser U, Waldschmidt D, et al. Gemcitabine plus erlotinib (GE) followed by capecitabine (C) versus capecitabine plus erlotinib (CE) followed by gemcitabine (G) in advanced pancreatic cancer (APC): a randomized, cross-over phase III trial of the Arbeitsgemeinschaft Internistische Onkologie (AIO). *J Clin Oncol* 2010; 28:(suppl 18S):abstract LBA4011.
51. Maurel J, Martin-Richard M, Conill C, et al. Phase I trial of gefitinib with concurrent radiotherapy and fixed 2-h gemcitabine infusion, in locally advanced pancreatic cancer. *Int J Radiat Oncol Biol Phys* 2006; 66: 1391–1398.
52. Fountzilas G, Bobos M, Kalogera-Fountzila A, et al. Gemcitabine combined with gefitinib in patients with inoperable or metastatic pancreatic cancer: a phase II study of the Hellenic Cooperative Oncology Group with biomarker evaluation. *Cancer Invest* 2008; 26: 784–793.
53. Schmidt-Ullrich RK, Mikkelsen RB, Dent P, et al. Radiation-induced proliferation of the human A431 squamous carcinoma cells is dependent on EGFR tyrosine phosphorylation. *Oncogene* 1997; 15: 1191–1197.
54. Dent P, Reardon DB, Park JS, et al. Radiation-induced release of transforming growth factor alpha activates the epidermal growth factor receptor and mitogen-activated protein kinase pathway in carcinoma cells, leading to increased proliferation and protection from radiation-induced cell death. *Mol Biol Cell* 1999; 10: 2493–2506.

55. Huang SM, Bock JM, Harari PM. Epidermal growth factor receptor blockade with C225 modulates proliferation, apoptosis, and radiosensitivity in squamous cell carcinomas of the head and neck. *Cancer Res* 1999; 59: 1935–1940.
56. Lammering G, Valerie K, Lin PS, et al. Radiosensitization of malignant glioma cells through overexpression of dominant-negative epidermal growth factor receptor. *Clin Cancer Res* 2001; 7: 682–690.
57. Krishnan S, Brown PD, Ballman KV, et al. Phase I trial of erlotinib with radiation therapy in patients with glioblastoma multiforme: results of North Central Cancer Treatment Group protocol N0177. *Int J Radiat Oncol Biol Phys* 2006; 65: 1192–1199.
58. Duffy A, Kortmansky J, Schwartz GK, et al. A phase I study of erlotinib in combination with gemcitabine and radiation in locally advanced, non-operable pancreatic adenocarcinoma. *Ann Oncol* 2008; 19: 86–91.
59. Munter M, Timke C, Abdollahi A, et al. Final results of a phase II trial [PARC-Study ISRCTN56652283] for patients with primary inoperable locally advanced pancreatic cancer combining intensity-modulated radiotherapy (IMRT) with cetuximab and gemcitabine. *J Clin Oncol* 2008; 26:(suppl):abstract 4613.
60. Borja-Cacho D, Jensen EH, Saluja AK, et al. Molecular targeted therapies for pancreatic cancer. *Am J Surg* 2008; 196: 430–441.
61. Xie K, Wei D, Huang S. Transcriptional anti-angiogenesis therapy of human pancreatic cancer. *Cytokine Growth Factor Rev* 2006; 17: 147–156.
62. Masood R, Cai J, Zheng T, et al. Vascular endothelial growth factor (VEGF) is an autocrine growth factor for VEGF receptor-positive human tumors. *Blood* 2001; 98: 1904–1913.
63. Seo Y, Baba H, Fukuda T, et al. High expression of vascular endothelial growth factor is associated with liver metastasis and a poor prognosis for patients with ductal pancreatic adenocarcinoma. *Cancer* 2000; 88: 2239–2245.
64. Kindler HL, Niedzwiecki D, Hollis D, et al. A double-blind placebo-controlled, randomized phase III trial of gemcitabine plus bevacizumab versus gemcitabine plus placebo in patients with advanced pancreatic cancer: a preliminary analysis of Cancer and Leukemia Group B (CALGB) 80303. Presented at: ASCO Gastrointestinal Cancers Symposium 2007; abstract 4508.
65. Fogelman DR, Varadhachary G, Xiong HQ, et al. Final results of a bi-institution phase II study of gemcitabine, oxaliplatin, and bevacizumab for advanced pancreatic cancer. Presented at: ASCO Gastrointestinal Cancers Symposium 2009; abstract 182.
66. Louvet C, Labianca R, Hammel P, et al. Gemcitabine in combination with oxaliplatin compared with gemcitabine alone in locally advanced or metastatic pancreatic cancer: results of a GERCOR and GISCAD phase III trial. *J Clin Oncol* 2005; 23: 3509–3516.
67. Kim GP, Oberg AL, Kabat B, et al. Phase II trial of bevacizumab, gemcitabine, and oxaliplatin in patients with metastatic pancreatic adenocarcinoma. Presented at: ASCO Gastrointestinal Cancers Symposium 2007; abstract 4113.
68. Javle M, Yu J, Garrett C, et al. Bevacizumab combined with gemcitabine and capecitabine for advanced pancreatic cancer: a phase II study. *Br J Cancer* 2009; 100: 1842–1845.
69. Siu LL, Awada A, Takimoto CH, et al. Phase I trial of sorafenib and gemcitabine in advanced solid tumors with an expanded cohort in advanced pancreatic cancer. *Clin Cancer Res* 2006; 12: 144–151.
70. Wallace JA, Locker G, Nattam S, et al. Sorafenib (s) plus gemcitabine (G) for advanced pancreatic cancer (PC): a phase II trial of the University of Chicago Phase II Consortium. *J Clin Oncol* 2007; 25:(suppl 18S):abstract 4608.
71. Spano JP, Chodkiewicz C, Maurel J, et al. Efficacy of gemcitabine plus axitinib compared with gemcitabine alone in patients with advanced pancreatic cancer: an open-label randomised phase II study. *Lancet* 2008; 371: 2101–2108.
72. Pfizer discontinues global phase III trial of axitinib for futility in advanced pancreatic cancer [press release]. New York, NY: Pfizer Inc.; January 30, 2009; http://mediaroom.pfizer.com/news/pfizer/20090202006193/en. Accessed on July 14, 2010.
73. Crane CH, Ellis LM, Abbruzzese JL, et al. Phase I trial evaluating the safety of bevacizumab with concurrent radiotherapy and capecitabine in locally advanced pancreatic cancer. *J Clin Oncol* 2006; 24: 1145–1151.
74. Crane CH, Winter K, Regine WF, et al. Phase II study of bevacizumab with concurrent capecitabine and radiation followed by maintenance gemcitabine and bevacizumab for locally advanced pancreatic cancer: Radiation Therapy Oncology Group RTOG 0411. *J Clin Oncol* 2009; 27: 4096–4102.
75. Vervenne WL, Bennouna J, Humblet Y, et al. A randomized, double-blind, placebo (P) controlled, multicenter phase III trial to evaluate the efficacy and safety of adding bevacizumab (B) to erlotinib (E) and gemcitabine (G) in patients (pts) with metastatic pancreatic cancer. *J Clin Oncol* 2008; 26: 4507–4509.
76. Von Hoff DD, Ramanathan R, Borad M, et al. SPARC correlation with response to gemcitabine (G) plus nab-paclitaxel (nab-P) in patients with advanced metastatic pancreatic cancer: a phase I/II study. *J Clin Oncol* 2009; 27:(suppl 15S):abstract 4525.
77. Hosein PJ, Pastorini VH, Gomez CM, et al. A phase II trial of nab-paclitaxel (NP) in patients with advanced pancreatic cancer (PC) who have progressed on gemcitabine-based therapy. Presented at: ASCO Gastrointestinal Cancers Symposium 2010; abstract 214.
78. Kayed H, Kleef J, Keleg S, et al. Indian hedgehog signaling pathway: expression and regulation in pancreatic cancer. *Int J Cancer* 2004; 110: 668–676.
79. Shafaee Z, Schmidt H, Du W, et al. Cyclopamine increases the cytotoxic effects of paclitaxel and radiation but not cisplatin and gemcitabine in Hedgehog expressing pancreatic cancer cells. *Cancer Chemother Pharmacol* 2006; 58: 765–770.
80. Feldmann G, Dhara S, Fendrich V, et al. Blockade of hedgehog signaling inhibits pancreatic cancer invasion and metastases: a new paradigm for combination therapy in solid cancers. *Cancer Res* 2007; 67: 2187–2196.
81. Hu WG, Liu T, Xiong JX, et al. Blockade of sonic hedgehog signal pathway enhances antiproliferative

effect of EGFR inhibitor in pancreatic cancer cells. *Acta Pharmacol Sin* 2007; 28: 1224–1230.

82. Ogino T, Onishi H, Suzuki H, Yamasaki A, Katano M. Evaluation of the hedgehog signaling pathway as a therapeutic target for dendritic cell-based immunotherapy against patients with cancer. *J Clin Oncol* 2010; 28:15-suppl, e13061.
83. Hakam A, Fang Q, Karl R, et al. Coexpression of IGF-1R and c-Src proteins in human pancreatic ductal adenocarcinoma. *Dig Dis Sci* 2003; 48: 1972–1978.
84. Kindler HL, Stephenson J, Garbo LE, et al. A placebo-controlled, randomized phase II study of conatumumab (C) or AMG 479 (A) or placebo (P) plus gemcitabine (G) in patients (pts) with metastatic pancreatic cancer (mPC). *J Clin Oncol* 2010; 28:(suppl 15S):abstract 4035.
85. Javle MM, Varadhachary GR, Bhosale P, et al. Phase I/II study of MK-0646, the humanized monoclonal IGF-1R antibody in combination with gemcitabine or gemcitabine plus erlotinib (E) for advanced pancreatic cancer. *J Clin Oncol* 2010; 28:(suppl):abstract 131.
86. Liu W, Bloom DA, Cance WG, et al. FAK and IGF-IR interact to provide survival signals in human pancreatic adenocarcinoma cells. *Carcinogenesis* 2008; 29: 1096–1107.
87. Ashworth A. Drug resistance caused by reversion mutation. *Cancer Res* 2008; 68: 10021–10023.
88. Jones S, Zhang X, Parsons DW, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* 2008; 321: 1801–1806.
89. Ischenko I, Camaj P, Seeliger H, et al. Inhibition of Src tyrosine kinase reverts chemoresistance toward 5-fluorouracil in human pancreatic carcinoma cells: an involvement of epidermal growth factor receptor signaling. *Oncogene* 2008; 27: 7212–7222.
90. Duxbury MS, Ito H, Zinner MJ, et al. RNA interference targeting the M2 subunit of ribonucleotide reductase enhances pancreatic adenocarcinoma chemosensitivity to gemcitabine. *Oncogene* 2004; 23: 1539–1548.
91. Hermann PC, Huber SL, Herrler T, et al. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell* 2007; 1: 313–323.
92. Ischenko I, Seeliger H, Kleespies A, et al. Pancreatic cancer stem cells: new understanding of tumorigenesis, clinical implications. *Langenbecks Arch Surg* 2010; 395: 1–10.
93. Mani SA, Guo W, Liao MJ, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008; 133: 704–715.
94. Pino MS, Balsamo M, Di Modugno F, et al. Human Mena+11a isoform serves as a marker of epithelial phenotype and sensitivity to epidermal growth factor receptor inhibition in human pancreatic cancer cell lines. *Clin Cancer Res* 2008; 14: 4943–4950.
95. Wolpin BM, Hezel AF, Ryan DP, et al. Phase II study of RAD001 in previously treated patients with metastatic pancreatic cancer. *J Clin Oncol* 2008; 26:(suppl 15S):abstract 4614.
96. Macdonald JS, McCoy S, Whitehead RP, et al. A phase II study of farnesyl transferase inhibitor R115777 in pancreatic cancer: a Southwest Oncology Group (SWOG 9924) study. *Invest New Drugs* 2005; 23: 485–487.
97. Evans JD, Stark A, Johnson CD, et al. A phase II trial of marimastat in advanced pancreatic cancer. *Br J Cancer* 2001; 85: 1865–1870.
98. Bramhall SR, Rosemurgy A, Brown PD, et al. Marimastat as first-line therapy for patients with unresectable pancreatic cancer: a randomized trial. *J Clin Oncol* 2001; 19: 3447–3455.
99. Bramhall SR, Schulz J, Nemunaitis J, et al. A double-blind placebo controlled, randomised study comparing gemcitabine and marimastat with gemcitabine and placebo as first line therapy in patients with advanced pancreatic cancer. *Br J Cancer* 2002; 87: 161–167.
100. Moore MJ, Hamm J, Dancey J, et al. Comparison of gemcitabine versus the matrix metalloproteinase inhibitor BAY 12-9566 in patients with advanced or metastatic adenocarcinoma of the pancreas: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* 2003; 21: 3296–3302.
101. Cohen SJ, Ho L, Ranganathan S, et al. Phase II and pharmacodynamic study of the farnesyltransferase inhibitor R115777 as initial therapy in patients with metastatic pancreatic adenocarcinoma. *J Clin Oncol* 2003; 21: 1301–1306.
102. Van Cutsem E, van De Velde H, Karasek P, et al. Phase III trial of gemcitabine plus tipifarnib compared with gemcitabine plus placebo in advanced pancreatic cancer. *J Clin Oncol* 2004; 22: 1430–1438.
103. Xiong HQ, Hess KR, Kayaleh OR, et al. A phase II trial of gemcitabine and celecoxib for metastatic pancreatic cancer. *J Clin Oncol* 2005; 23:(suppl):abstract 4174.
104. Ferrari V, Valcamonico F, Amoroso V, et al. Gemcitabine plus celecoxib (GECO) in advanced pancreatic cancer: a phase II trial. *Cancer Chemother Pharmacol* 2006; 57: 185–190.
105. Dragovich T, Burris H, Loehrer P, et al. Gemcitabine plus celecoxib in patients with advanced or metastatic pancreatic adenocarcinoma: results of a phase II trial. *Am J Clin Oncol* 2008; 31: 157–162.
106. Lipton A, Campbell-Baird, Cynthia RN, et al. Phase II trial of gemcitabine, irinotecan, and celecoxib in patients with advanced pancreatic cancer. *J Clin Gastroenterol* 2010; 44: 286–288.
107. Ko AH, Dicke K, Gurtler J, et al. Phase II, randomized, open-label study of cetuximab and bevacizumab alone or in combination with fixed-dose rate (FDR) gemcitabine as first-line therapy for patients with metastatic adenocarcinoma of the pancreas (MPC). Presented at: ASCO Gastrointestinal Cancers Symposium 2009; abstract 183.
108. Kindler HL, Karrison T, Hochster HS, et al. Final analysis of a randomized phase II study of bevacizumab (B) and gemcitabine (G) plus cetuximab (C) or erlotinib (E) in patients (pts) with advanced pancreatic cancer (PC). *J Clin Oncol* 2008; 26:(suppl):abstract 4502.

Melanoma Immunotherapy

Shanthi Sivendran, MD, Bradley Glodny, MD, Michael Pan, Miriam Merad, MD, PhD, and Yvonne Saenger, MD

Mount Sinai School of Medicine, New York, NY

OUTLINE

CYTOKINES
- **Interleukin-2**
- **Interleukin-21**
- **Interferon-alfa**
- **Granulocyte-Macrophage Colony-Stimulating Factor**
- **Tumor Necrosis Factor**

IMMUNOREGULATORY ANTIBODIES
- **Anti–Cytotoxic T-Lymphocyte Antigen-4**
- **Anti-Programmed-1**
- **Anti-CD137**
- **Anti-CD40**

ADOPTIVE T-CELL THERAPY

MELANOMA VACCINES
- **Whole-Cell Vaccines**
- **Peptide/Protein-Based Vaccines**
- **Ganglioside Vaccine**
- **Dendritic Cell–Based Therapy**
- **Recombinant Viral Vectors and DNA Vaccines**

CONCLUSION

ABSTRACT

Melanoma immunotherapy has been an area of intense research for decades, and this work is now yielding more tangible results for patients. Work has focused on 4 main areas: cytokine therapy, administration of immune-modulating antibodies, adoptive T-cell therapy, and vaccines. Cytokine therapy is an established treatment for advanced melanoma, and immune-modulating antibodies have recently emerged as an exciting new area of drug development with efficacy now established in a phase III trial. Adoptive T-cell therapy provides the proof of principle that T cells can attack and eliminate tumors. It has been challenging, however, to adapt this treatment for widespread use. Vaccines have generally yielded poor results, but intratumor pathogen-based strategies have shown encouraging results in recent trials, perhaps due to stronger immune stimulation. A review of the field of melanoma immunotherapy is provided here, with emphasis on those agents that have reached clinical testing. Novel strategies to induce the immune system to attack melanomas are reviewed. In the future, it is envisioned that immunotherapy will have further application in combination with cytotoxic and targeted therapies. ***Mt Sinai J Med 77:620–642, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: antibody, cytokine, immunotherapy, melanoma, vaccine.

The treatment of advanced melanoma has generally been disappointing. The only 2 therapies approved by the US Food and Drug Administration (FDA) for the treatment of advanced melanoma are dacarbazine and interleukin 2 (IL-2), with response rates ranging from 5% to 20%.[1] Given that the median survival for advanced melanoma is 6 months, alternative therapies that are either curative in intent or disease stabilizing are desperately needed.[1] Despite this need, few therapies have translated into improved overall survival. A meta-analysis of all phase II cooperative group trials in stage IV melanoma demonstrated a median survival time of 6.2 months, with 25.5% of patients alive at 1 year.[2] A review of 35 immunotherapy trials involving 765 patients demonstrated an overall response rate of 3.3%.[3]

The potential role of the immune system in the treatment of cancer was initially proposed over the past 100 years from the use of bacterial toxins in the treatment of cancer to the observation of spontaneous remissions in melanoma.[4] More recently, immunological advances such as anti-cytotoxic T-lymphocyte antigen-4 (anti-CTLA-4), the first therapeutic, immune-based, anticancer vaccine have been

Address Correspondence to:

Yvonne Saenger
Mount Sinai School of Medicine
New York, NY
Email: yvonne.saenger@exchange.mssm.edu

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20215

© 2010 Mount Sinai School of Medicine

developed. This is now the first therapy to show a survival benefit in a large randomized phase III trial in advanced melanoma. As our understanding of immunoregulation becomes increasingly sophisticated, the ability to overcome tumor evasion of antitumor immune responses and tumor-induced immune suppression has increased. New therapies targeting immunoregulatory mechanisms and novel vaccination strategies have been developed with encouraging results, generating new optimism in the field of melanoma immunotherapy. In this review we discuss some of the most promising immunotherapies that have translated from the preclinical arena to clinical trials.

CYTOKINES

Interleukin-2

First identified in 1976, IL-2 is the first and only FDA-approved immunotherapy for melanoma at the present time. From 1985 to 1994, a total of 270 patients with metastatic melanoma received high-dose therapy consisting of 720,000 IU/kg of IL-2 administered intravenously every 8 hours for a maximum of 15 doses per cycle. Seventeen patients with metastatic melanoma had complete regression of all disease and 26 patients had partial regression. This data established high-dose bolus IL-2 as an effective treatment for advanced melanoma. Toxicity has always been a very significant problem with high-dose IL-2 therapy, which induces a syndrome resembling sepsis. The most common adverse effects were hypotension (64%), with a few patients experiencing grade IV effects (1%); and supraventricular tachycardia (17%). Six patients died from treatment-related toxicity.[5,6]

> *First identified in 1976, interleukin-2 is the first and only FDA-approved immunotherapy for melanoma at the present time.*

Due to the high toxicity of IL-2, mainly on hemodynamic instability, research has focused on identifying which patients are likely to sustain benefit, and also on the development of more tolerable regimens. Notably, the extent of metastases affects patient outcomes. In a retrospective study, 371 patients with metastatic melanoma who received high-dose IL-2 were analyzed. Patients with metastases only in cutaneous tissues had a higher response rate (50%) compared with patients who had metastases at visceral sites (13%) and at both sites (14%). This suggests that patients with lesions limited to the cutaneous and subcutaneous level could be better candidates for high-dose IL-2 therapy.[7] Lower dosages were tested to minimize toxicity. In a phase II trial, 25 patients with renal cell carcinoma, malignant melanoma, and colorectal cancer were given low-dose IL-2 subcutaneously in a cycle of 5 days of treatment followed by 9 days of rest. This treatment was well tolerated, with no grade II adverse effects. However, although the response rate was 33.3% in patients with renal cell carcinoma, no major responses were observed in patients with malignant melanoma or colorectal cancer.[8]

Targeting the tumor site is another strategy to minimize toxicity while maintaining patient response. In a recent phase II study, 51 patients with metastatic melanoma received intratumoral IL-2 injections 3 times weekly. Complete responses that lasted >6 months were observed in 70% of injected lesions, but no responses in visceral disease were observed. Only grade I and II toxicity was observed, mostly at the site of injection.[9] A review of these trials is presented in Table I.

Interleukin-21

Interleukin-21 regulates the life cycle of B cells, playing a role in both production and apoptosis in certain settings.[10] Interleukin-21 can also promote activation of CD8+ T cells and NK cells, increasing their cytotoxicity.[11] Several recent trials have studied the use of IL-21 in patients with metastatic melanoma. In a phase II trial, 28 patients with unresectable metastatic melanoma received a daily intravenous bolus injection of recombinant IL-21 at 30 μg/kg/dose for 5 days, followed by 9 days of rest. The cycle continued for 6 weeks, followed by 2 weeks without treatment. Complete and partial responses were achieved in 1 patient each. The treatment was well tolerated with primarily grade I and II adverse events.[12] In another recent phase II study, 22 patients with chemotherapy-naïve metastatic or recurrent melanoma were assigned different dosing regimens of intravenous IL-21. Five patients had partial responses and 9 patients had stable disease. Adverse effects were mild, with grade I and II fatigue and muscle pain being the most common.[13] This study is ongoing. Further trials will need to be conducted to clarify the potential of IL-21 as a useful treatment against melanoma.

Interferon-alfa

Interferon-alfa (IFNa) is released by lymphocytes and is observed to have antiviral, antiproliferative,

Table 1. *Summary of Cytokine-Related Trials.*

Cytokine	Author	Type	N	Disease	Efficacy	Side Effects	Dosage
IL-2	Rosenberg *et al.*[6]	Composite	283	Metastatic melanoma or renal cancer	17 complete regression, 26 partial regression	Severe; hypotension (64%) and supraventricular tachycardia (17%) in all grades. Nausea, vomiting, and diarrhea were common.	720,000 IU/kg intravenously every 8 hours for a maximum of 15 doses per cycle. Treatment course consisted of 2 cycles.
IL-2	Tagliaferri *et al.*[8]	Phase II	25	Renal cell carcinoma, malignant melanoma, colorectal cancer	33% response rate in renal cell carcinoma patients; no major responses in melanoma or colorectal cancer	Well-tolerated; no grade II adverse effects	1 mIU rIL-2 subcutaneous injection every 8 hours for 5 days, followed by 9 days of rest.
IL-2	Weide *et al.*[9]	Phase II	51	Metastatic melanoma	70% of injected lesions had complete response for >6 months; 69% of patients achieved complete local response of all treated metastases	Well-tolerated; mostly grade I and II toxicity at site of injection	Intratumoral injection of 0.3 mIU to 6 mIU IL-2, depending on lesion size. Up to 25 lesions treated simultaneously. Treatment given 3× weekly.
IL-21	Davis *et al.*[12]	Phase II	28	Unresectable metastatic melanoma	1 CR and 1 PR observed; overall response rate 8.3%	Well-tolerated; mostly grade I and II flu-like illness	Intravenous bolus injection of rIL-21 at 30 μg/kg/dose daily for 5 days, followed by 9 days of rest. Cycle continued for 6 weeks, followed by 2 weeks without treatment.
IL-21	Petrella *et al.*[13]	Phase II (ongoing)	22	Chemotherapy-naïve metastatic or recurrent melanoma	5 PR and 9 SD	Well-tolerated; mostly grade I and II fatigue and muscle pain	Either 50 μg or 30 μg/kg/day intravenous bolus injection for 5 days every other week for 2 to 6 weeks
IFNa	Kirkwood *et al.* (E1684)[14]	Composite	287	Melanoma with lesions at Breslow depth > 4 mm; all patients received regional lymph node dissection	Increase in DFS (1.72 versus 1.00 years); no significant difference in OS	67% grade II and 9% grade IV effects	20 mU/m^2/day, 5 days/week for 4 weeks, then 3 times weekly at 10 mU/m^2/day for 48 weeks, or close observation

(*continued overleaf*)

Table 1. *(Continued)*.

Cytokine	Author	Type	N	Disease	Efficacy	Side Effects	Dosage
IFNa	Kirkwood *et al.* (E1690)[14]	Composite	642	Similar criteria to E1684 except regional lymph node dissection not mandated	No difference in OS between high-dose IFNa and observation	Mostly grade III toxicity; no toxicity-related deaths	High-dose (20 mU/m^2/day IV, 5 days/week for 4 weeks) or low-dose IFNa-2b (3 mU/m^2/day), or observation. Dosage reduced for toxicity.
IFNa	Hauschild *et al.*[15]	Phase III	850	Resected cutaneous melanoma	No difference in OS or DFS between 18- and 60-month treatment groups	Well-tolerated; adverse effects mild to moderate. Significantly more adverse effects in 60-month treatment group.	3 mU injections 3× weekly for 18 or 60 months
GM-CSF	Spitler *et al.*[19]	Phase II	48	Melanoma rendered clinically disease free after surgery	Increase in median survival (37.5 months versus 12.2 months) against matched historical controls	Well-tolerated; mild adverse effects including transient myalgias, fatigue, and rashes	Daily 125 μg/m^2 injections for 14 days, followed by 14 days of rest. Cycle continued for 1 year or until disease recurrence.
GM-CSF	Spitler *et al.*[20]	Phase II	98	Stage II, III, or IV melanoma		Well-tolerated; 82% of patients experienced grade I and II adverse effects	Same treatment as previous trial except cycle continued for 3 years
GM-CSF	Lawson *et al.*[21]	Phase III	815	Completely resected stage III or IV melanoma	No significant difference in OS; median OS 72,1 months for treatment arm versus 59.8 months for placebo arm	Well-tolerated; toxicity consistent with other trials	250 μg or placebo daily for 14 days, followed by 14 days of rest. Cycle continued for up to 1 year.
TNF	Rossi *et al.*[23]	Phase I	59	Melanoma with in-transit metastases	Complete response rate 63% for melphalan and TNF versus 41.5% for melphalan only	Mild to moderate; no difference between treatment groups	Isolated limb perfusion of melphalan with or without low-dose TNF.

Abbreviations: CR, complete response; DFS, disease-free survival; GM-CSF, granulocyte-macrophage colony-stimulating factor; IFNa, interferon-alfa; IL-2, interleukin-2; IL-21, interleukin-21; OS, overall survival; PR, partial response; rIL-2, recombinant interleukin-2; SD, stable disease; TNF, tumor necrosis factor.

and immunomodulatory effects. High-dose IFNa has been studied in several trials conducted by the Eastern Cooperative Oncology Group. In trial E1684, 287 "high-risk" melanoma patients with a Breslow depth >4 mm received either high-dose IFNa-2b or close observation. The results showed an increase in disease-free survival (DFS) (1.72 versus 1.00 years) and in overall survival (OS) (3.82 versus 2.8 years) in the treatment arm. Subsequent analysis, however, unfortunately showed that the increase in OS was no longer significant. Toxicity was considerable, with 67% of patients experiencing grade II adverse effects and 9% experiencing grade IV. It is important to note that 2 patients died from hepatotoxicity related to treatment.[14] Trial E1694, comparing high-dose IFNa with a GM2-KLH vaccine (a vaccine consisting of GM2 ganglioside, a melanoma-specific antigen, conjugated with the immunostimulant keyhole limpet hemocyanin), was closed prematurely due to improved survival in the IFNa arm. The meaning of these results was subsequently called into question, unfortunately, when a subsequent trial suggested that the control vaccine was inferior to observation.[14]

Lower dosages were tested to see if toxicity could be reduced while maintaining patient response. Using similar eligibility requirements as E1684, trial E1690 enrolled 642 patients to receive either high-dose IFNa, low-dose IFNa, or observation. After a follow-up of 52 months, no difference in OS was noted.[14] Studies of lower-dose interferon remain inconclusive, with none conclusively proving equivalence or inferiority to the high-dose regimen, and none showing improvement in OS. Prolonged treatment was tested to see whether patient response could be improved. In a phase III trial, 850 patients with resected cutaneous melanoma were given 3 mU IFNa injections subcutaneously for either 18 months or 60 months. Most patients were lymph node–negative and received sentinel node biopsy. Prolonged treatment did not significantly improve patient response; no difference was observed in DFS or OS.[15] In summary, trials of IFNa have consistently shown an improvement in relapse-free survival (RFS) but not OS, with some suggestion that survival benefit may end with the discontinuation of treatment. Therapy is associated with significant decrement in quality of life for many patients. Subset analysis has suggested that OS benefit may be achieved in patients with ulcerated primary lesions, but this has not been validated in randomized prospective studies.[16]

Trials of interferon-alfa have consistently shown an improvement in relapse-free survival but not overall survival. Subset analysis has suggested that overall survival benefit may be achieved in patients with ulcerated primary lesions, but this has not been validated in randomized prospective studies.

Granulocyte-Macrophage Colony-Stimulating Factor

Granulocyte-macrophage colony-stimulating factor (GM-CSF) is important for the activation of macrophages and the recruitment of peripheral blood monocytes to become cytotoxic to tumor cells.[17] It can also activate dendritic cells (DCs), increasing antigen presentation necessary for a T cell–mediated immune response.[18] Granulocyte-macrophage colony-stimulating factor has been used in patients with varying results. In a phase II trial, 48 melanoma patients who were rendered clinically disease free after surgery were treated with GM-CSF daily at 125 $\mu g/m^2$ for 14 days, followed by 14 days of rest. This treatment cycle continued for 1 year or until disease recurrence. Median survival was 37.5 months for the treatment group versus 12.2 months for matched historical controls. The survival rate was 89% at 1 year and 64% at 2 years. The treatment was well tolerated, with mild adverse effects including transient myalgias, fatigue, and rashes.[19] The effect of prolonged treatment was subsequently tested. The same 28-day treatment cycle of GM-CSF was given for 3 years to 98 patients with stage II, III, or IV melanoma. The 5-year melanoma survival rate was 60%. The treatment was generally well tolerated, with grade I or II side effects occurring in 82% of patients. It is important to note that 2 patients undergoing 3 years of treatment developed acute myeloid leukemia.[20]

In E4697, a large-scale phase III group study of 815 patients with completely resected stage III or IV melanoma was conducted. Patients received 250 mcg GM-CSF or placebo daily for 14 days every 28 days, for up to 1 year. Overall survival did not achieve statistical significance, with median OS at 72.1 months for the treatment arm versus 59.8 months for the placebo arm. Toxicity was consistent with other trials.[21] Subset analysis suggested possible selective benefit for patients with stage IV disease, but this has yet to be confirmed in a prospective randomized trial.[19]

Tumor Necrosis Factor

Tumor necrosis factor (TNF) is produced by macrophages and was discovered in 1975 for its ability to induce hemorrhagic necrosis of tumor cells.[22] Clinical application has been limited due to significant toxicity. A study was conducted to compare patient response to melphalan with or without TNF. Fifty-nine patients with in-transit metastases received isolated limb perfusion of both melphalan and low-dose TNF versus 53 in a control group who only received melphalan. The complete response rate was 63% for patients treated with TNF and 41.5% for patients receiving only melphalan. Toxicity was mild to moderate, with no difference noted between treatment groups. The results from this study demonstrated that low-dose TNF may be useful for in-transit disease.[23]

IMMUNOREGULATORY ANTIBODIES

Anti–Cytotoxic T-Lymphocyte Antigen-4

During T-cell activation, cluster of differentiation 28 (CD28) binds with B7 costimulatory molecules expressed on mature antigen-presenting cells (APCs). This binding promotes cytokine production and affects downstream targets that induce cell-cycle progression. During this activation, CTLA-4 (an inhibitory coreceptor) is up-regulated and competes with CD28 for binding to B7. Ligated CTLA-4 sends an inhibitory downstream signal to block cytokine production and cell-cycle progression.[24] This was demonstrated in early preclinical models in which mice depleted of CTLA-4 demonstrated massive lymphoproliferation leading to destruction of major organs from lymphocytic infiltration.[25–27] These studies prompted further research into inhibiting the B7/CTLA-4 interaction to enhance T-cell activation. In primary prostate cancer mouse models, treatment with antibodies to CTLA-4 reduced tumor incidence significantly when combined with an irradiated tumor-cell vaccine; and in a metastatic prostate cancer mouse model, the use of CTLA-4 blockade after primary prostate tumor resection significantly reduced metastatic relapse.[28,29] Combination therapy of CTLA-4 blockade and a GM-CSF vaccine eradicated tumors in 80% of B16 mouse melanoma models.[30] Given the strength of the preclinical data, 2 monoclonal antibodies to CTLA-4 were developed for human clinical trials: ipilimumab and tremelimumab. A summary of all immunoregulatory antibody–related clinical trials is presented in Table 2.

Ipilimumab

Several phase I studies have suggested a role for CTLA-4 blockade in melanoma while maintaining a reasonable safety profile. A pilot study of 14 patients with metastatic melanoma was treated with intravenous anti–CTLA-4 antibody (Ab) at 3 mg/kg followed by subcutaneous vaccination with 2 modified human leukocyte antigen A (HLA-A)*0201-restricted peptides from the glycoprotein 100 (gp100) melanoma-associated antigen, gp100:209-217(210M) and gp100:280-288(288V). Two complete responses and 1 partial response were observed, with 6 patients experiencing grade III and IV autoimmune manifestations: dermatitis, enterocolitis, hepatitis, and hypophysitis.[31] This same trial was then expanded to 56 patients with 2 different dosing schedules. Twenty-nine patients received 3 mg/kg of anti–CTLA-4 every 3 weeks and 27 patients received this same initial dose followed by reduced doses of 1 mg/kg every 3 weeks. Both groups continued to receive concomitant vaccination with 2 modified HLA-A*0201-restricted peptides from the gp100 melanoma-associated antigen, gp100:209-217(210M) and gp100:280-288(288V). Thirteen percent of patients achieved an overall objective response rate, with 2 complete responses and 5 partial responses. Grade III and IV autoimmune toxicity–colitis, dermatitis, uveitis, enterocolitis, hepatitis, and hypophysitis–was experienced by 36% of patients. There was no significant difference in response rate or toxicity between the 2 groups.[32]

A phase II study of 217 patients with pretreated advanced melanoma (stage III and IV) randomly assigned patients to either 10 mg/kg, 3 mg/kg, or 0.3 mg/kg of ipilimumab every 3 weeks for 4 cycles, followed by maintenance therapy with their assigned ipilimumab dose every 12 weeks until progression, death, or toxic effects requiring discontinuation. Patients receiving 10 mg/kg had the best overall response (11.1%), followed by patients dosed at 3 mg/kg (4.2%). Patients who received 0.3 mg/kg had no response. Grade III/IV adverse events (diarrhea) occurred in 11 out of 73 of the patients receiving the 10 mg/kg dosing schedule.[33] Another phase II study of pretreated, unresectable stage III/IV melanoma showed an overall response rate of 5.8% and a disease-control rate of 27% when patients were administered 10 mg/kg of ipilimumab every 3 weeks for 4 cycles, followed by maintenance therapy every 3 months. Median overall survival in this group was 10.2 months.[34] Given the significant diarrhea patients may experience with ipilimumab therapy, a phase II trial of patients with unresectable advanced-stage (III or IV) melanoma were randomized to receive ipilimumab (10 mg/kg every 3 weeks for 4

Table 2. *Immunoregulatory Antibody Trials.*

Author	Phase	Drug	Disease	N	Efficacy				Side Effects	Dosage
					CR	PR	SD	Description		
Phan *et al.*[31]	I	Ipilimumab	Metastatic melanoma	14 in initial trial	2	1	NA		6 patients experienced grade III/IV autoimmune manifestations. Symptoms included dermatitis, enterocolitis, hepatitis, and hypophysitis.	3 mg/kg intravenously followed by subcutaneous vaccination with 2 modified HLA-A* 0201-restricted peptides from gp100:209–217 (210M) and gp100:280-288(288V)
Attia *et al.*[32]	I	Ipilimumab	Pretreated advanced melanoma	56	2	5	NA		14 patients experienced grade III/IV autoimmune toxicity. Symptoms included dermatitis, enterocolitis, hepatitis, and hypophysitis.	2 patients received 3 mg/kg intravenously every 3 weeks and 27 patients received 3 mg/kg intravenously initially and then dose reduction to 1 mg/kg IV every 3 weeks. Both groups also had sc vaccination with 2 modified HLA-A* 0201-restricted peptides from gp100:209-217(210M) and gp100:280–288 (288V)
Wolchok *et al.*[33]	II	Ipilimumab	Pretreated advanced melanoma	217	2 0 0	6 3 0	13 16 10	Best overall response was 11.1% in the 10 mg/kg dosing group	Immune-related events occurred in 50 of 71 patients in the 10 mg/kg dosing group. The most common grade III/IV events were gastrointestinal.	Patients were assigned to either: 10 mg/kg intravenously every 3 weeks for 4 cycles (n = 73), or 3 mg/kg intravenously every 3 weeks for 4 cycles (n = 72), or 0.3 mg/kg intravenously every 3 weeks for 4 cycles (n = 72)
O'Day *et al.*[34]	II	Ipilimumab	Pretreated advanced melanoma	5	0	5	33		19% grade III and 3.2% grade IV immune-related skin and gastrointestinal events	10 mg/kg intravenously every 3 weeks for 4 cycles (induction) followed by maintenance therapy every 3 months

(*continued overleaf*)

Table 2. *(Continued).*

Author	Phase	Drug	Disease	N	Efficacy					Side Effects	Dosage
					CR	PR	SD	Description			
Weber *et al.*[35]	II	Ipilimumab	Pretreated and treatment-naïve patients with advanced melanoma	115	1 0	6 9	11 11	Budesonide should not be used as prophylaxis for diarrhea ≥ grade II		Diarrhea ≥ grade II occurred in 32.7% of patients treated with budesonide and 35% of those treated with placebo.	Patients were assigned to either: 10 mg/kg intravenously every 3 weeks for 4 doses with daily budesonide, or 10 mg/kg intravenously every 3 weeks for 4 cycles with placebo
Hodi *et al.*[36]	III	Ipilimumab	Pretreated advanced melanoma	676	1 2 0	22 13 2	58 24 13	Ipilimumab with or without ap100 peptide vaccine compared with gp100 alone improved OS (10 months)		60% of patients treated with ipilimumab had immune-related events; the frequency of grade III/IV events was 10%–15%. The skin and GI tract (diarrhea) were the most common sites affected. Of the 2.1% of deaths related to study drugs, half were from immune-related adverse events.	Patients were assigned to either: 3 mg/kg intravenously every 3 weeks for up to 4 doses plus gp100, or 3 mg/kg intravenously every 3 weeks for up to 4 doses alone, or gp100 alone
Ribas *et al.*[37]	I	Tremelimumab	Advanced solid malignancies	39	2	2	4	10 mg/kg was considered the MTD		The most common dose-limiting toxicities were dermatitis and diarrhea.	Dose-escalation study of 7 doses starting at 0.01 mg/kg
Camacho *et al.*[38]	I/II	Tremelimumab	Metastatic melanoma	28 89	1	3	NA	No dose-limiting toxicity in phase I 10% attained objective responses		13% grade III/IV adverse events in 15 mg/kg arm and 27% in the 10 mg/kg once every month. Most frequent events were diarrhea, rash, and pruritis	Monthly intravenous infusions of 3, 6, or 10 mg/kg for up to 1 year to determine recommended dose in phase I portion, or 10 mg/kg once every month, or 15 mg/kg every 3 months

(continued overleaf)

Table 2. *(Continued)*.

Author	Phase	Drug	Disease	N	Efficacy				Side Effects	Dosage
					CR	PR	SD	Description		
Kirkwood *et al.*[39]	II	Tremelimumab	Refractory/relapsed melanoma	251	0	16	35	Objective response rate of 6.6%, with all responses being durable >170 days	Grade III/IV adverse events included diarrhea (11%), fatigue (2%), and colitis (4%). There were 2 treatment-related deaths.	15 mg/kg every 90 days
Ribas *et al.*[40]	III	Tremilimumab	Unresectable advanced melanoma	655	NA	NA	NA	Awaiting publication		Patients were randomized to either 15 mg/kg tremelimumab intravenously every 3 months, or TMZ 200 mg/m^2 orally, or DTIC 1000 mg/m^2
Brahmer *et al.*[46]	I	Anti-PDCD1	Refractory advanced solid tumors	39	1	2	NA		1 serious adverse event (inflammatory colitis) in a patient who received 5 doses of 1 mg/kg	Single intravenous infusion of anti-PD-1 in dose-escalating 6-patient cohorts at 0.3, 1, 3, or 10 mg/kg, followed by a 15-patient expansion cohort at 10 mg/kg
Sznol *et al.*[47]	I	Anti-PDCD1	Refractory metastatic NSCLC, RCC, MEL, or CRPC	16 in each disease category	NA	NA	NA	Still enrolling		Dose-escalating study of 1, 3, or 10 mg/kg intravenously. Enrollment at the MTD was expanded to 16 patients in each disease category.

(continued overleaf)

Table 2. *(Continued).*

					Efficacy					
Author	Phase	Drug	Disease	N	CR	PR	SD	Description	Side Effects	Dosage
Brahmer *et al.*[48]	II	Anti-PDCD1	Refractory metastatic NSCLC, RCC, MEL, CRPC, or CC	21	NA	1	NA	Awaiting publication. 2 mixed responses in MEL patients		10 mg/kg intravenously in single dose with follow-up at 8 weeks. Patients with SD or intralesional response could receive additional treatment at weeks 12 and 16.
Sznol *et al.*[54]	I	Anti-CD137	MEL, RCC, or ovarian cancer	83	0	3	4	Awaiting publication	Most common grade III/IV events: fatigue (3%), reversible transaminitis (11%), and neutropenia (5%)	Part 1: dose-escalating study. Part 2: randomized to 1, 3, and 10 mg/kg
Khalil *et al.*[55]	I	CD40 agonist mono-clonal antibody	Advanced solid tumors	29	0	4	7	Well-tolerated across all doses	Most common adverse event was grade I/II cytokine release syndrome, which included chills, rigors, and fever	Dose-escalation study from 0.01 to 0.3 mg/kg
Khalil *et al.*[55]	I	Recombinant human CD40 ligand	Advanced solid tumors and NHL	32	0	2	12		Grade III/IV transaminase elevations occurred in up to 57% of patients (receiving 0.15 mg/kg/day dose)	Dose-escalation study

Abbreviations: CC, colon cancer; CR, complete response; CRPC, castrate-resistant prostate cancer; DTIC, dacarbazine; HLA-A, human leukocyte antigen A; MEL, melanoma; MTD, maximally tolerated dose; NA, not applicable; NHL, non-Hodgkin lymphoma; NSCLC, non-small-cell lung cancer; OS, overall survival; PDCD1, programmed cell-death 1; PR, partial response; RCC, renal cell carcinoma; SD, stable disease; TMZ, temozolamide.

doses) with daily budesonide or the same dosing of ipilimumab with a daily placebo. Budesonide did not affect the rate of grade II to IV diarrhea. Patients who experienced grade III/IV adverse effects had a higher disease-control rate than those who experienced grade 0 to II adverse effects.[35]

Recently a phase III study using ipilimumab in 676 patients with stage II or IV melanoma showed improvement in median OS (10.1 months) compared with gp100 alone (6.4 months). The dosing schema for ipilimumab in this study was 3 mg/kg every 3 weeks for up to 4 treatments. Grade III/IV immune-related events occurred in 10%–15% of patients treated with ipilimumab, and 14 deaths were associated with the study drug. The most common adverse effect was diarrhea.[36] This is the first trial to show a significant difference between treatment arms in the metastatic-disease setting. A trial of ipilimumab in combination with dacarbazine versus dacarbazine alone as front-line treatment was recently completed, but results have not been reported to date.

> *Recently a phase III study using ipilimumab in 676 patients with stage II or IV melanoma showed improvement in median OS (10.1 months) compared with gp100 alone (6.4 months). This is the first trial to show a significant difference between treatment arms in the metastatic-disease setting.*

Tremelimumab

Tremelimumab (CP-675,206) has been investigated in phase I, II, and III trials. A phase I trial of 39 patients with a variety of advanced solid-tumor malignancies demonstrated that this drug can be administered safely as a single dose up to 15 mg/kg, with 2 patients experiencing complete responses, 2 with partial responses, and 4 with stable disease. Dose-limiting toxicities included diarrhea, dermatitis, vitiligo, panhypopituitarism, and hyperthyroidism.[37] Another phase I/II trial first demonstrated safety and tolerability of tremelimumab at a dose of up to 10 mg/kg in 28 patients with metastatic melanoma. During the phase II portion, the 10 mg/kg dose was administered monthly or 15 mg/kg every 3 months (based on previous trials). Ten percent of patients had objective tumor responses, and there was 1 complete response and 3 partial responses in each dosing regimen. Responses were durable up to 30 months. Grade III/IV adverse events affected 13% of patients in the 15 mg/kg arm and 27% in the 10 mg/kg arm, with diarrhea, rash, and pruritus being the most common events.[38] Another phase II trial of refractory/relapsed melanoma looked at 15 mg/kg of tremelimumab every 3 months for up to 4 doses. In patients with tumor response or stable disease, 4 additional doses could be given. An objective response rate of 6.6% was demonstrated, with a durable response rate of >170 days. Eleven percent of patients had grade III/IV events, including diarrhea, fatigue, and colitis, and there were 2 treatment-related deaths.[39] A phase III trial in patients with advanced melanoma compared tremelimumab (15 mg/kg every 3 months) to temozolomide or dacarbazine. As a single agent, tremelimumab showed no difference in overall survival compared with chemotherapy (11.1 months versus 10.7 months).[40]

Anti-Programmed-1

Programmed cell death (PD-1) belongs to the CD28 family of coreceptors expressed on activated T and B cells and binds to 2 ligands, PD-L1 and PD-L2. Ligand PD-L1 is expressed on T cells, B cells, macrophages, and DCs. Ligand PD-L2 is expressed on DCs and activated macrophages.[41] A marker of T-cell exhaustion, PD-1 associated with immunosuppression, and PD-1 knockout mice develop autoimmune-like diseases such as lupus-like disease and autoimmune dilated cardiomyopathy.[42,43] Tumors and the tumor microenvironment also express PD-L1, and blockading this may be a strategy for immunotherapy. Murine myeloma cell lines naturally express PD-L1. Administering a PD-1 Ab to normal mice infected with these myeloma lines demonstrated significant suppression in the myeloma growth. This same experiment in PD-1–deficient mice led to complete suppression of tumor growth.[44] In human ovarian cancer, higher expression of PD-L1 was associated with a statistically significant poorer outcome than lower expression. There was also a significant inverse correlation between this increased expression of PD-L1 and decreased anti-tumor CD8+ T cells.[41] PD-1–deficient mice injected with B16 melanoma cells in the flank region had increased T-cell response and cytokine production, with subsequent inhibition of hematogenous spread of melanoma cells to the liver.[45] Based on these preclinical findings, a phase I trial of a single agent anti-PD 1 (MDX-1106) Ab was administered to treatment-refractory solid-tumor patients at a dose of either 0.3, 1, 3, or 10 mg/kg. Among 39 patients, 1 had a complete response and 2 had partial responses. One serious adverse event (inflammatory colitis) was observed at a dose of 1 mg/kg.[46] Currently, a phase

I trial of MDX-1106 in advanced refractory solid-tumor malignancies is under way with escalating biweekly doses of 1, 3, or 10 mg/kg.[47] A second phase II trial to evaluate the safety of a single infusion of 10 mg/kg in refractory/relapsed malignancies has demonstrated clinical activity in melanoma without serious toxicity.[48] A combination trial targeting CTLA-4 (ipilimumab) and PD-1 (MDX-1106) in melanoma, as well as a combination trial of PD-1 (MDX 1106) and peptide vaccines comprising gp100:209-217(210M), MART-1:26-35(27L), gp100:280-288(288V), and NY-ESO-1(165V) emulsified in Montanide ISA 51 VG, is currently recruiting patients.[49,50]

Anti-CD137

Part of the TNF receptor superfamily that delivers costimulatory signals to T cells, 4-1BB is expressed on activated immune cells (T cells, natural killer cells, DCs, eosinophils, mast cells).[51] Binding of 4-1BB to its ligand up-regulates antiapoptotic genes and prevents T-cell death.[51] Mice deficient in either the receptor or ligand have decreased ability to respond to infections such as influenza and lymphocytic choriomeningitis.[51] Antibodies targeting 4-1BB potentially could increase T-cell response and possibly T-cell memory. A preclinical study combining 4-1BB Ab with GM-CSF–secreting tumor cell immunotherapy in B16 melanoma model mice demonstrated rejection of established tumors.[52] Another study combining anti-4-1BB with depletion of CD4 showed increased immune infiltration of immune cells into tumors, reduction of regulatory barriers, and increased differentiation of CD8+ T cells.[53] Based on results such as these, a phase I study of human anti-CD137 (BMS-663513) was conducted in patients with advanced solid-tumor malignancies. It was administered intravenously every 3 weeks with a dose escalation. This drug was found to be safe across a dosing range of 0.3–15 mg/kg, with the primary grade II/IV events being fatigue, reversible transaminitis, and neutropenia.[54]

Anti-CD40

Also a member of the TNF receptor superfamily, CD40 is a transmembrane protein that is involved in both immune activation and tumor apoptosis.[55] It is normally expressed on macrophages, DCs, B cells, platelets, and other nonimmune cells, with a corresponding ligand expressed on activated T cells and platelets. Signaling via CD40-stimulated inflammatory cytokines activates T cells and enhances the expression of major histocompatibility complex (MHC) and costimulatory molecules. Those who have mutations in either CD40 or its ligand are immunosuppressed and unable to mount an adequate immune response to infection.[56–58] Interestingly, ligation of CD40 on tumor cells results in apoptosis secondary to a direct cytotoxic effect.[59,60] Based on this, agonist Abs to CD40 have entered clinical development. Three phase I trials using a formulation of CD40 (recombinant human trimer, full human immunoglobin G [IgG]2 mAb, or humanized IgG1 mAb) have been published. All of these combinations are agonists of CD40, although the humanized IgG1 mAb is a weak agonist. Patients in 2 of the trials were of mixed advanced solid tumors and in 1 trial were either advanced, refractory non-Hodgkin lymphoma or multiple myeloma. Six partial responses were seen in the solid-tumor trials, and 3 partial responses and 2 complete remissions were seen among the non-Hodgkin lymphoma patients. The main toxicities were a cytokine-release syndrome, increased aspartate aminotransferase/alanine aminotransferase, and injection-site reactions.

ADOPTIVE T-CELL THERAPY

Adoptive cell therapy exploits melanoma's unique ability to produce large amounts of antitumor lymphocytes in vivo. These antitumor lymphocytes are harvested from melanoma tissue and then grown ex vivo. The ability to grow these cells ex vivo allows for significant numbers of antitumor lymphocytes and the ability to select cells with specific tumor antigens. Potentially immunosuppressive cells can be removed. Clinically, this technique has been optimized by the

Adoptive cell therapy exploits melanoma's unique ability to produce large amounts of antitumor lymphocytes in vivo. These antitumor lymphocytes are harvested from melanoma tissue and then grown ex vivo in significant numbers. Objective response rates to these regimens in selected patients with metastatic melanoma have reached up to 72% (49%– 72%), with durable responses in complete responders for up to 63 months.

introduction of lymphodepletion with chemoradiation.[61] This may improve the host environment by eliminating competition for homeostatic cytokines and by removing T regulatory cells. Objective response rates to these regimens in selected patients with metastatic melanoma have reached up to 72% (49%–72%), with durable responses in complete responders for up to 63 months.[62] These response rates are very exciting, but limiting factors include the expense of the therapy and the fact that tumor-infiltrating lymphocytes cannot be cultured from all patients.

MELANOMA VACCINES

The first cancer vaccine was developed serendipitously in 1890 by William Coley, who administered bacterial toxins to cancer patients with advanced disease.[4] Since this time, promising results have been presented within the area of cancer vaccines. During the 1970s, early melanoma vaccine research included that by Donald Morton utilizing the Bacillus Calmette-Guérin (BCG) vaccine, a form of nonspecific active immunotherapy.[63] Most recently, the FDA approved the first cancer vaccine, Provenge, for the treatment of prostate cancer using autologous DC precursors pulsed with the prostatic acid phosphatase peptide.

> *The FDA recently approved the first cancer vaccine, Provenge, for the treatment of prostate cancer using autologous dendritic cell precursors pulsed with the prostatic acid phosphatase peptide.*

Over the past 10 to 15 years, tumor immunotherapy has been a major area of research in the field of melanoma. During this time, much work has been performed to assess the immunogenicity of melanoma antigens including melanosomal proteins, cancer-testis antigens, and mutated gene products.[64] Many antigens have been found to be potential immunologic targets eliciting both humoral and cellular immune responses. Professional APC's (ie, DCs) play a large role in priming specific cytotoxic T-cell and B-cell responses. However, these responses are often not strong enough to eradicate melanoma, possibly indicating tolerance induction.[65,66] This lack in response is likely due to the immunosuppressive microenvironment produced by the tumor.[67] Therefore, a strategy to effectively and consistently produce a durable boost in the immune system's response to these antigens is integral to the efficacy of melanoma vaccines in the therapy of stage III/IV melanoma.[66,68]

Thus far, melanoma vaccines have taken a variety of shapes and forms, including whole-cell tumor preparations, cell lysates, peptide/protein, DC, naked DNA, gangliosides, heat shock proteins, and recombinant viral vectors, in addition to several other newer vaccines still in the early phases of development. Although phase I/II trials have been promising, phase III/IV trials for the most part have yielded disappointing results. As seen in Table 3, phase III trial results have been found to be either insignificant in terms of clinical benefit or even possibly detrimental. This disconnect between phase I/II and phase III/IV trials is quite worrisome and evidence of the need for a better way to evaluate these vaccines in early trials. For the most part, phase I/II trials measure efficacy indirectly via in vitro T-cell assays such as enzyme-linked immunosorbent spot (ELISPOT) or tetramer staining and often possess no control arm. Additionally, phase III trials in the adjuvant setting take an extremely long time to obtain results and generate data, often >10 years, as recurrences may be delayed. This is not to say, however, that a melanoma vaccine cannot be effective. In light of such failures, many researchers have begun to focus on combining therapies and considering a number of adjuvants to help strengthen the immune system's response.[69]

Whole-Cell Vaccines

Whole-cell melanoma vaccines are theorized to hold potential for melanoma therapy due to the common antigens found on the cells. Therefore, these cells, either autologous or allogenic, can be irradiated/inactivated and administered to induce an immune response toward the common antigens on the melanoma cells. Early work by Stein *et al.* utilized an autologous irradiated whole-cell melanoma vaccine combined with BCG in the treatment of both breast cancer and melanoma in the late 1970's.[70] Additionally, Hollinshead *et al.* employed an allogenic irradiated whole-cell lysate with promising results in the early 80's.[71] Many phase I/II clinical trials have found promising results; however, phase III/IV trials have fallen short of these expectations. One of the most thoroughly studied is the CanVaxin vaccine, a polyvalent, antigen-rich, allogeneic whole-cell vaccine containing 3 melanoma cell lines, expressing >20 antigenic tumor proteins covering >95% of HLA phenotypes.[65,72] In the largest phase III vaccine trial (n = 1656) including stage III/IV melanoma, BCG alone was compared with BCG plus CanVaxin. This trial was terminated prematurely due to a lack of improvement in survival relative to BCG alone.[65,72,73]

Table 3. *Completed Phase III Melanoma Vaccine Trials.*

Investigator	Stage	Vaccine	Control	N	Vaccine		Control	
					DFS	OS	DFS	OS
Whole-cell vaccines								
Morton *et al.*[73]	III	Canvaxin + BCG	Placebo + BCG	1160	47.2%	59.1%	52.1%	67.7%
	IV (resected)	Canvaxin + BCG	Placebo + BCG	496	27.4%	39.6%	20.9%	44.9%
Sondak *et al.*/SWOG[75]	T3N0M0	Melacine	Observation	689	65%	NA	63%	NA
Viral vector vaccines								
Mitchell *et al.*[74]	III	Melacine + LD IFN	HD IFN	604	NA	9.4 mo	NA	12.3 mo
Hersey *et al.*[109]	IIb/III	VMCL	Observation	700	6.98 years	>8.45 years	4.37 years	7.34 years
Peptide/protein vaccines								
Schwartzentruber *et al.*[90]	III/IV	gp100:209-217(210M) + Montanide ISA + HD IL-2	HD IL-2	185	2.9 months	17.6 months	1.6 months	12.8%
O'Day *et al.*[91]	III/IV	gp100 + ipilimumab or either alone	Ipilimumab	676	2.8 months	10.0 months	2.9 months	10.1 months
MEDAREX	III/IV	MDX-1379 + MDX-010	MDX-010	750	75.1%	89.2%	77.8%	92.4%
Ganglioside vaccines[91]								
Livingston *et al.*[96]	IIb/III	GM2 + BCG	BCG	123	48%	57%	30%	46%
Kirkwood *et al.*[97]	III	GM2-KLH/QS-21	IFNa-2b	880	49%	73%	62%	78%
Eggermont *et al.*[99]	T3-T4N0M0	GM2-KLH/QS-21	Observation	1350	68.2%	81.5%	69.4%	83.6%
HSP vaccines								
Testori *et al.*[95]	IV	Autologous HSP gp96 complexes	Physician's choice	322	NS*	NA	NA	NA
DC vaccines								
Schadendorf *et al.*[104]	IV	DC	DTIC	108	NA	26.4%	NA	31.0%

Abbreviations: BCG, Bacillus Calmette-Guérin; DC, dendritic cell; DFS, disease-free survival; DTIC, dacarbazine; ECOG, Eastern Cooperative Oncology Group; GM2-KLH; vaccine consisting of GM2 ganglioside conjugated with the immunostimulant keyhole limpet hemocyanin; HD, high-dose; HSP, heat shock protein; IFN, interferon; IFNa, interferon-alfa; IL-2, interleukin-2; LD, low-dose; NA, not applicable; NS, not significant; OS, overall survival; SWOG, Southwest Oncology Group; VMCL, vaccinia melanoma cell lysate; VMO, vaccinia melanoma oncolysate.

*Overall, no significance was found; however, upon substage analysis (M1a and M1b) of patients who received ≥10 vaccines, improved OS compared with control group was found.

Other studies utilizing the Melacine vaccine, comprising 2 allogenic cell line lysates combined with the adjuvant DETOX (detoxified Freund's adjuvant), demonstrated similar yet somewhat more promising results.[65,74] In Southwest Oncology Group's phase III randomized observation-controlled Melacine trial of stage IIa melanoma, there was no significant benefit in RFS compared with the observation arm (hazard ratio: 0.92, $P = 0.51$).[75] However, upon post-hoc analysis there appeared to be an association in RFS and 5 HLA class I haplotypes. In response to these findings, Sosman *et al.* performed class I (HLA-A, B, and C) serotyping on 553 (80%) of the 689 patients. A significant difference in 5-year RFS (83% versus 59%; $P = 0.0005$) was found in HLA-A2 and C3 patients (58%) compared with the control arm.[65,76,77] A possible explanation for this association is the presence of class I antigens in the vaccine lysate that is being presented in the context of A2 and C3.[77] This relative failure of allogenic whole-cell and lysate vaccines in phase III trials may be due to the lack of epitope/protein congruity between the allogenic vaccines with the specific tumor of the patient. This possibility is the basis for the development of autologous irradiated whole-cell vaccines cultured from the patient's tumor.[78] In phase I/II studies, Mvax was found to elicit an antitumor response assessed by delayed-type hypersensitivity (DTH) and at the highest doses elicit no serious toxicity. Major downsides of these autologous vaccines do exist, however; primarily, the cost and time to produce such vaccines from patient tumors are restricting and have limited the role of such vaccines.[78,79]

Although phase III studies of allogeneic/autologous vaccines have been disappointing, attention has also been focused on adjuvants used with these whole-cell vaccines in order to boost the immune response. There are a multitude of adjuvants being developed for use with all types of melanoma immunotherapy, but specific to whole-cell vaccines are the intratumoral gene transfer modified cell lines. These cell lines are engineered to secrete a cytokine such as GM-CSF, FLT-3L, or IL-2, helping to enhance the host immune response. The best studied among these is the GM-CSF–secreting whole-cell vaccine, likely because mouse experiments showed efficacy for GM-CSF–secreting melanoma cells in the prophylactic setting.[79–81] The mechanism of action is hypothesized to be the ability of GM-CSF to activate, multiply, and recruit professional APCs, which are thought to present tumor antigens to effector T cells.[79] Early phase I trials have been promising, particularly by Soiffer *et al.* They transduced autologous melanoma cells via an adenoviral-mediated gene transfer to express GM-CSF. Seventeen out of 25 (68%) patients exhibited DTH responses as well as DC, macrophage, and lymphocyte infiltration in 19/26 (73%) injection sites.[64] There were 2 regressions in tumor, 1 complete and 1 partial.[64,65]

Peptide/Protein-Based Vaccines

In 1991, MAGE-1 was the first tumor peptide cloned.[4,82] Since then, hundreds of tumor peptides, lysates, and protein antigens have been made accessible for melanoma peptide vaccination. Of these, many have been used in vaccines, including MAGE-3, NY-ESO, MART-1, tyrosinase, gp100, and survivin. However, these peptides are self-epitopes restricted by MHC class I molecules, and a way in which to "break peripheral tolerance" is a necessity. Early phase I/II studies vaccinating with 1 peptide/protein alone have failed to induce specific T-cell responses in patients.[83,84] Therefore, several techniques have been developed to help stimulate a stronger and more durable immune response. These include the use of multiple peptides in a single vaccine, and the addition of immune stimulatory adjuvants such as helper peptides, molecules to help induce the MHC-I CD4 T-cell (TH2)–mediated response,[85] and cytokines to help break tolerance.[86] Although these peptide vaccines plus adjuvant cocktails have been found to often elicit strong cytotoxic T-cell responses measured by tetramers and ELISPOT, rarely has tumor regression coincided with this response.[82] However, in few instances, antigen-specific T-cell responses induced by peptide vaccines have been associated with dramatic clinical responses.[66]

Phase I/II trials vaccinating patients with multiple protein/peptides have been promising in light of peptide vaccines' low toxicity profile and evidence of cytotoxic T cell (CTL) peptide specific activation in in-vitro assays.[83,87,88] Additionally, a phase II study administered a gp100:209-2176(210M) vaccine with high-dose IL-2. The researchers found a clinical response rate of 42% (13/31), plus an additional 4 patients who experienced mixed or minor responses.[89] However, no larger phase III trials have been able to replicate such a high response.

Promising results were seen in a prospective randomized phase III trial of 185 HLA-201 melanoma patients with stage III/IV disease. Arm 1 ($n = 94$) received HD IL-2(720,000 IU/kg/dose) alone, and arm 2 ($n = 91$) received gp100:209-217(210M) peptide vaccine + Montanide ISA followed by HD IL-2.[90] The response rate and progression-free survival both were found to be significantly higher in arm 2 than in arm 1. Response rate was 22.1% versus 9.7%, respectively (by chi-square test, $P = 0.0223$) and progression-free survival was 2.9 months (range, 1.7–4.5 months) versus 1.6 months (range, 1.5–1.8 months), respectively, $P = 0.0101$. This study is the first phase III trial to show a significant benefit of both peptide/protein vaccines and melanoma vaccines in general.[90]

In contrast, a recent phase III trial by O'Day *et al.* assessed the efficacy of anti–CTLA-4 (ipilimumab) versus the same gp100 vaccine (peptides 209-217[210M] and 280-288[288V]) alone and in combination in HLA-A*0201$^+$ on unresectable stage III/IV melanoma (n = 676). Anti–CTLA-4 was found to be statistically superior to gp100 vaccine in disease-control rate, best overall response rate, and OS. Additionally, the gp100 vaccine did not add any significant benefit in terms of risk reduction (32%–34%) or any of the outcome endpoints when given in combination with anti–CTLA-4.[91] Furthermore, 32 of these patients who had progressed following an initial response to the induction course of therapy were reinduced (anti–CTLA-4-8, combination-23, gp100-1) as per the initial protocol.[92] They found the disease-control rate in the anti–CTLA-4 arm to be 65%–75% compared with 0% in the gp100 arm after reinduction. Whereas these results speak strongly for the use of anti–CTLA-4 in the treatment of metastatic melanoma, they are disappointing in terms of peptide vaccines. The reasons why a peptide vaccine would add benefit to IL-2 but not to ipilimumab are unclear. These studies suggest the need for immune-modulating adjuvants such as Montanide ISA, although the efficacy gap could be related to differences in the currently incompletely understood mechanisms of action of high-dose IL-2 and anti–CTLA-4 mAb.

Additional peptide targets under vaccine development are tumor-derived heat shock protein-peptide complexes (HSPPC). Heat shock proteins are highly conserved, chaperone proteins involved in intracellular transport and protein turnover and have been found to simulate the innate and adaptive arms of the immune system.[93,94] HSPPC-96 or vitespen (formerly Oncophage), an autologous vaccine derived from patient's tumors, has shown promising results in phase I/II trials and has recently completed phase III trials in melanoma and renal cancer.[93] In the melanoma phase III trial, the HSPPC-96 vaccine efficacy was compared with physician's choice (regimens containing IL-2 and/or dacarbazine and/or complete surgical resection) in stage IV melanoma patients.[93] The study enrolled 322 patients, of whom 215 were randomly assigned into the vaccine arm and 107 into the physician's choice arm. Patients in the vaccine arm had surgery to isolate the tumor for their vaccine, and in about 6 to 8 weeks they received 25-mg doses of HSPPC-96 subcutaneously for 4 weeks and every other week subsequently. Of the 215 patients in the vaccine arm, 82 patients (38.1%) did not receive the vaccine. Of these patients, 61/82 (74%) did not receive the vaccine due to technical difficulties associated with vaccine preparation, therefore highlighting a major obstacle of autologous-derived vaccines.[95] Total OS was found to be similar between both arms, even when stratified into M1a, M1b, and M1c substages. However, in those patients who received 10 or more vaccines, a clear benefit was noted in all substages analyzed together, in M1a and M1b separately, but not in M1c. Therefore, this suggests that HSPPC-96 may be best utilized in a specific subpopulation of patients and as an adjuvant in those patients who can receive 10 or more vaccines.[94,95]

Ganglioside Vaccine

Antigens other than protein/peptides have also been used to create melanoma vaccines. For example, gangliosides are N-linked complex sialoglycosphingolipids and are found on the surface of melanoma cells as well as several other types of cancers. Although they are not peptide/protein vaccines, they have often been conjugated with proteins to increase their immunogenicity. In early phase I/II trials, the GM2 ganglioside vaccine was found to be immunogenic in some patients, stimulating the production of antiganglioside Abs which were cytotoxic to melanoma cells expressing GM2.[96,97] Additionally, those patients which produced these Abs proved to experience significantly greater disease-free interval and DFS.[98] The most widely studied has been the GM2-KLH/QS-21 vaccine. In the largest phase III metastatic melanoma trial to date (n = 1314), the GM2-KLH/QS-21 vaccine was compared with observation in stage II (T3-T4N0M0) melanoma. The study was ended prematurely, as the observation arm was shown to have a superior relapse-free survival compared with the vaccine arm. Therefore, not only is the GM2-KLH/QS-21 vaccine ineffective, it may even be potentially harmful. This detrimental effect is thought to be due to the induction of tolerance induced by the vaccine.[99]

Dendritic Cell–Based Therapy

The DC is the most potent APC. It was first discovered in the spleen by its unparalleled ability to stimulate antigen-specific T cells.[76] Therefore, DC-based therapy has been in development as a use for advanced melanoma therapy for many years. In studies such as that by Abdel-Wahab *et al.*, it was shown that DCs pulsed with soluble peptide or tumor lysates were able to induce a specific CD8 CTL response.[100] Unfortunately, autologous DCs of cancer patients have been found to be unresponsive or defective.[76] Consequently manipulation of these DCs ex vivo and more recently in vivo is necessary for their activation.[76] This manipulation often includes stimulation of DCs with particular melanoma antigens, including peptides, proteins, tumor lysates, and messenger RNA (mRNA) or DNA.[76,101]

Initial phase I/II studies have shown both evidence of immune reactivity as well as clinical response.[102,103] Several multicenter randomized phase III/IV studies are currently under way attempting to elucidate the clinical benefit of such vaccines. In a large phase III trial (n = 108) examining the efficacy of autologous peptide-loaded DCs, the vaccine failed to improve disease burden or survival versus dacarbazine in stage IV melanoma patients, closing early due to a clear lack of efficacy.[104] Although these results were disappointing, this trial did bring issues to light regarding DC vaccine development, including information on route of administration (intradermal > subcutaneous), consequence of HLA haplotype on clinical response, ex vivo stimulation of DCs, and selection of adjuvants to further optimize DC vaccination.[104]

Several more phase I/II studies have been completed or are currently under way, most of which showing activation of the adaptive immune response, with some even showing prolonged survival and long-lived antigen-specific immune memory induced by these DC vaccines.[105] These trials have utilized DCs pulsed/loaded with a variety of tumor antigens, including mRNA-encoding tumor antigens, melanoma-derived peptides, and tumor lysates, in addition to the use of different adjuvants.[105–107] However, large phase III/IV trials have failed to replicate these results, highlighting the fact that immune activation does not equal clinical efficacy.

Dendritic cells have been proven time and time again to stimulate directed adaptive T-cell responses; however, they lack proven clinical efficacy in larger phase III/IV trials to date. With the development of more immunogenic adjuvants, improved stimulation of DCs, and more efficient routes of administration, there will likely be future advancements in the development of DC-based vaccines.

Recombinant Viral Vectors and DNA Vaccines

Viral Vectors

Several viral vectors have been used in the development of recombinant viral vaccines, including adenovirus, poxvirus, vaccinia, and herpes simplex virus 1 (HSV-1).[76,108,109] Several clinical trials have been performed to date inducing vectors to express specific tumor antigens, cytokines, or a number of different costimulatory molecules (B7.1, ICAM-1, LFA-3).[65,76] Therefore, it is possible to create a viral vector with several modes to stimulate the immune response. This method of vaccination is advantageous compared with other forms because an entire melanoma antigen can be transduced into a viral vector, therefore activating a broader range of T-cell–specific responses.[76] One significant obstacle in using viral vectors, however, is the immunogenicity of the viral particles themselves. These viral antigens are often stronger activators of the human immune system compared with the melanoma antigens. Therefore, in many patients, neutralizing Abs against the vaccine are produced.[65,76] Currently studies are under way utilizing other less-immunogenic vectors for the vaccine, such as the fowlpox virus.[110]

Initially, results from phase I/II trials were disappointing; however, with the addition of costimulatory molecules such as B7.1, more promising results were attained.[65,111] More recently, Kaufman *et al.* completed a phase I trial combining 3 costimulatory molecules (B7.1, ICAM-1, and LFA-3) called TRICOM transduced into a vaccinia vector. Although it was proven to be safe and potentially effective (30.7% objective clinical response), larger trials are currently under way to further characterize the role of antigens, adjuvant cytokines, and costimulators with the vaccinia TRICOM vaccine.[112]

Of late, Wallack *et al.* tabulated the 10-year analysis of a phase III trial comparing the efficacy of vaccinia melanoma oncolysate (VMO) vaccine versus attenuated vaccinia virus (VV; negative control).[113] Patient data were collected from 11 institutions between June 1988 and June 1991. At time of analysis, information on 109 patients was available, with 78 living (35 VMO, 43 VV). The results were somewhat disappointing, showing no difference between patients who received the VMO and those who received the attenuated VV. However, some promise does exist, as statistical significance was noted in males aged 44–57 years; in addition, the fact that 71.5% (78/109) of patients are still alive today indicates some activation of the immune system.[113]

Recently, the HSV-1 vector has been used in phase I/II trials with some promising results. OncoVEX GM-CSF(JS1/34.5-/47-/GM-CSF) represents a newer form of HSV-1 vaccine that exhibits direct oncolytic properties to tumor and the development of a systemic immune response.[108] This second-generation HSV-I vector has been engineered to increase oncolysis in tumor cells and to express

GM-CSF in order to increase recruitment of DCs.[108,114] In a recent phase II trial, Senzer *et al.* treated 50 patients with stage IIc/IV metastatic melanoma and found a response rate of 26% by Response Evaluation Criteria in Solid Tumors (RECIST) and overall survival at 1 year to be 54%.[115] Based upon these results, a phase III clinical trial is currently under way.[110]

> *In a recent phase II trial, a herpes simplex virus-1 vaccine that exhibits direct oncolytic properties to tumor and the development of a systemic immune response was used to treat 50 patients with stage IIc/IV metastatic melanoma. The trial showed a response rate based on Response Evaluation Criteria in Solid Tumors (RECIST) of 26%, and an overall survival at 1 year of 54%.*

DNA

Another approach in developing a melanoma vaccine is the use of DNA-encoding melanoma antigens. It has long been understood that DNA can induce a systemic immune response, activating both humoral and cellular immunity.[76,116,117] Advantages of this approach include the lack of a viral vector, hence no viral antigens to induce a strong immune response against the vaccine itself. Additionally, bacterial DNA contains a natural immune adjuvant in the form of CpG, a potent immune stimulator.[76] Another advantage to this form of immunization is that HLA haplotype is not of concern and therefore may be useful in a broad patient population.[65]

Providing evidence for the use of plasmid DNA is Tagawa *et al.* in a phase I trial that administered DNA plasmids encoding tyrosinase epitopes to stage IV melanoma patients. They found both a specific T-cell response (11/26) as well as an increase in survival (16/26) in a majority of these patients.[118] However, it has proven challenging to administer large doses of DNA to humans; to date, there have been no large phase III/IV trials.

CONCLUSION

Successful management of melanoma with immunotherapy depends not only on our understanding of tumor biology, but also on the host immunological milieu. Overcoming tumor-evasive and -suppressive mechanisms, as well as gaining a clearer understanding of immunoregulatory mechanisms, is crucial for therapeutic drug development. Although earlier therapies in melanoma have had disappointing results, the production of more targeted therapies has translated to response rates and in some cases improved OS in the metastatic setting.

Several important issues should be addressed in the use of immunotherapy. First, careful patient selection is important for the success of several of these therapies. Interleukin-2 has demonstrated complete remissions in a select population. Given its significant toxicity, it is only appropriate for younger patients with minimal comorbidities and adequate organ function. Although not demonstrated in clinical trials, INFa may have some benefit in the adjuvant setting, specifically in patients with ulcerated lesions. Adoptive cell therapy has had very impressive response rates in the metastatic setting but may also favor a patient population that is healthier and able to withstand lymphodepletion and radiation therapy. Identifying which patients would most benefit from such treatments is important in providing patient-tailored care.

Second, targeting melanoma through multiple pathways may be more successful than single-agent therapy. In particular, with the development of novel targeted agents against the proto-oncogene B-Raf, combining immunomodulatory treatments with B-Raf inhibition and/or cytotoxic therapy is exciting and has the potential to yield significant benefit. Combining multiple immunotherapeutic modalities may also be effective. Vaccine therapies in particular, although limited in success, prime the immune system and promote a systemic immune response. In combination with other therapies, such as IL-2, there has been an improvement in tumor response rates. This suggests a role in combination therapy. It is important to note that this strategy is not always successful, as in the case of combination ipilimumab and the peptide vaccines. Clearly, a greater understanding of how each therapy modulates the immune is needed. Currently, several clinical trials are recruiting patients for combination immune therapy, and the results of these will lead to future insights.

Although ideally the intent for all these therapies is curative, the prognosis for advanced melanoma is so poor that disease stabilization with prolonged survival and improved quality of life are also important endpoints. With the development of immunoregulatory Abs such as anti-CTLA-4, some of these endpoints are being realized. Anti–CTLA-4 has also raised challenges regarding the assessment of patient response, as in some cases survival is improved although objective response is not achieved by standard criteria. Excitingly, ipilimumab is the first of these immunotherapy agents to show an improvement in OS in the phase III setting. Although early in clinical trials, other immunomodulatory Abs also show significant promise. Vaccines including strong adjuvants, such as HSV vectors, have also yielded encouraging results.

Advanced melanoma is a devastating disease, but with our improved understanding of tumor immunology and new therapies in the pipeline, as well as the development of newer targeted therapies, we are beginning to alter the disease course of advanced melanoma. A brighter future is ahead.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Eggermont AM, Schadendorf D. Melanoma and immunotherapy. *Hematol Oncol Clin North Am* 2009; 23: 547–564, ix–x.
2. Korn EL, Liu PY, Lee SJ, et al. Meta-analysis of phase II cooperative group trials in metastatic stage IV melanoma to determine progression-free and overall survival benchmarks for future phase II trials. *J Clin Oncol* 2008; 26: 527–534.
3. Rosenberg SA, Yang JC, Restifo NP. Cancer immunotherapy: moving beyond current vaccines. *Nat Med* 2004; 10: 909–915.
4. Kirkwood JM, Tarhini AA, Panelli MC, et al. Next generation of immunotherapy for melanoma. *J Clin Oncol* 2008; 26: 3445–3455.
5. Atkins MB, Lotze MT, Dutcher JP, et al. High-dose recombinant interleukin 2 therapy for patients with metastatic melanoma: analysis of 270 patients treated between 1985 and 1993. *J Clin Oncol* 1999; 17: 2105–2116.
6. Rosenberg SA, Yang JC, Topalian SL, et al. Treatment of 283 consecutive patients with metastatic melanoma or renal cell cancer using high-dose bolus interleukin 2. *JAMA* 1994; 271: 907–913.
7. Chang E, Rosenberg SA. Patients with melanoma metastases at cutaneous and subcutaneous sites are highly susceptible to interleukin-2-based therapy. *J Immunother* 2001; 24: 88–90.
8. Tagliaferri P, Barile C, Caraglia M, et al. Daily low-dose subcutaneous recombinant interleukin-2 by alternate weekly administration: antitumor activity and immunomodulatory effects. *Am J Clin Oncol* 1998; 21: 48–53.
9. Weide B, Derhovanessian E, Pflugfelder A, et al. High response rate after intratumoral treatment with interleukin-2: results from a phase 2 study in 51 patients with metastasized melanoma. *Cancer* 2010; 116: 4139–4146.
10. Ozaki K, Spolski R, Ettinger R, et al. Regulation of B cell differentiation and plasma cell generation by IL-21, a novel inducer of Blimp-1 and Bcl-6. *J Immunol* 2004; 173: 5361–5371.
11. Frederiksen KS, Lundsgaard D, Freeman JA, et al. IL-21 induces in vivo immune activation of NK cells and CD8(+) T cells in patients with metastatic melanoma and renal cell carcinoma. *Cancer Immunol Immunother* 2008; 57: 1439–1449.
12. Davis ID, Brady B, Kefford RF, et al. Clinical and biological efficacy of recombinant human interleukin-21 in patients with stage IV malignant melanoma without prior treatment: a phase IIa trial. *Clin Cancer Res* 2009; 15: 2123–2129.
13. Petrella T, Savage K, Tozer R, et al. Phase II study of interleukin-21 in patients with metastatic or recurrent melanoma. Abstract presented at: Seventh World Melanoma Congress; May 12–16, 2009; Vienna, Austria.
14. Kirkwood JM, Manola J, Ibrahim J, et al. A pooled analysis of Eastern Cooperative Oncology Group and intergroup trials of adjuvant high-dose interferon for melanoma. *Clin Cancer Res* 2004; 10: 1670–1677.
15. Hauschild A, Weichenthal M, Rass K, et al. Efficacy of low-dose interferon {alpha}2a 18 versus 60 months of treatment in patients with primary melanoma of ≥1.5 mm tumor thickness: results of a randomized phase III DeCOG trial. *J Clin Oncol* 2010; 28: 841–846.
16. Eggermont AM, Suciu S, Testori A, et al. Ulceration of primary melanoma and responsiveness to adjuvant interferon therapy: analysis of the adjuvant trials EORTC18952 and EORTC18991 in 2644 patients. *J Clin Oncol* 2009; 27(suppl 15S): abstract 9007.
17. Grabstein KH, Urdal DL, Tushinski RJ, et al. Induction of macrophage tumoricidal activity by granulocyte-macrophage colony-stimulating factor. *Science* 1986; 232: 506–508.
18. Szabolcs P, Moore MA, Young JW. Expansion of immunostimulatory dendritic cells among the myeloid progeny of human CD34+ bone marrow precursors cultured with c-kit ligand, granulocyte-macrophage colony-stimulating factor, and TNF-alpha. *J Immunol* 1995; 154: 5851–5861.
19. Spitler LE, Grossbard ML, Ernstoff MS, et al. Adjuvant therapy of stage III and IV malignant melanoma using granulocyte-macrophage colony-stimulating factor. *J Clin Oncol* 2000; 18: 1614–1621.
20. Spitler LE, Weber RW, Allen RE, et al. Recombinant human granulocyte-macrophage colony-stimulating factor (GM-CSF, sargramostim) administered for 3 years as adjuvant therapy of stages II(T4), III, and IV melanoma. *J Immunother* 2009; 32: 632–637.
21. Lawson DH, Lee SJ, Tarhini AA, et al. Phase III cooperative group study of yeast-derived granulocyte macrophage colony-stimulating factor (GM-CSF) versus placebo as adjuvant treatment of patients with completely resected stage III-IV melanoma. *J Clin Oncol* 2010; 28(suppl 15S): abstract 8504.
22. Carswell EA, Old LJ, Kassel RL, et al. An endotoxin-induced serum factor that causes necrosis of tumors. *Proc Natl Acad Sci U S A* 1975; 72: 3666–3670.
23. Rossi CR, Pasquali S, Mocellin S, et al. Long-term results of melphalan-based isolated limb perfusion with or without low-dose TNF for in-transit melanoma metastases. *Ann Surg Oncol* 2010;: doi:10.1245/s10434-010-1104-2.

24. Krummel MF, Allison JP. CD28 and CTLA-4 have opposing effects on the response of T cells to stimulation. *J Exp Med* 1995; 182: 459–465.
25. Khattri R, Auger JA, Griffin MD, et al. Lymphoproliferative disorder in CTLA-4 knockout mice is characterized by CD28-regulated activation of Th2 responses. *J Immunol* 1999; 162: 5784–5791.
26. Tivol EA, Borriello F, Schweitzer AN, et al. Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity* 1995; 3: 541–547.
27. Waterhouse P, Penninger JM, Timms E, et al. Lymphoproliferative disorders with early lethality in mice deficient in CTLA-4. *Science* 1995; 270: 985–988.
28. Kwon ED, Foster BA, Hurwitz AA, et al. Elimination of residual metastatic prostate cancer after surgery and adjunctive cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) blockade immunotherapy. *Proc Natl Acad Sci U S A* 1999; 96: 15074–15079.
29. Hurwitz AA, Foster BA, Kwon ED, et al. Combination immunotherapy of primary prostate cancer in a transgenic mouse model using CTLA-4 blockade. *Cancer Res* 2000; 60: 2444–2448.
30. Van Elsas A, Hurwitz AA, Allison JP. Combination immunotherapy of B16 melanoma using anti-cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) and granulocyte/macrophage colony-stimulating factor (GM-CSF)-producing vaccines induces rejection of subcutaneous and metastatic tumors accompanied by autoimmune depigmentation. *J Exp Med* 1999; 190: 355–366.
31. Phan GQ, Yang JC, Sherry RM, et al. Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma. *Proc Natl Acad Sci U S A* 2003; 100: 8372–8377.
32. Attia P, Phan GQ, Maker AV, et al. Autoimmunity correlates with tumor regression in patients with metastatic melanoma treated with anti-cytotoxic T-lymphocyte antigen-4. *J Clin Oncol* 2005; 23: 6043–6053.
33. Wolchok JD, Neyns B, Linette G, et al. Ipilimumab monotherapy in patients with pretreated advanced melanoma: a randomised, double-blind, multicentre, phase 2, dose-ranging study. *Lancet Oncol* 2010; 11: 155–164.
34. O'Day SJ, Maio M, Chiarion-Sileni V, et al. Efficacy and safety of ipilimumab monotherapy in patients with pretreated advanced melanoma: a multicenter single-arm phase II study. *Ann Oncol* 2010; 21: 1712–1717.
35. Weber J, Thompson JA, Hamid O, et al. A randomized, double-blind, placebo-controlled, phase II study comparing the tolerability and efficacy of ipilimumab administered with or without prophylactic budesonide in patients with unresectable stage III or IV melanoma. *Clin Cancer Res* 2009; 15: 5591–5598.
36. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 2010; 363: 711–723.
37. Ribas A, Camacho LH, Lopez-Berestein G, et al. Antitumor activity in melanoma and anti-self responses in a phase I trial with the anti-cytotoxic T lymphocyte-associated antigen 4 monoclonal antibody CP-675,206. *J Clin Oncol* 2005; 23: 8968–8977.
38. Camacho LH, Antonia S, Sosman J, et al. Phase I/II trial of tremelimumab in patients with metastatic melanoma. *J Clin Oncol* 2009; 27: 1075–1081.
39. Kirkwood JM, Lorigan P, Hersey P, et al. Phase II trial of tremelimumab (CP-675,206) in patients with advanced refractory or relapsed melanoma. *Clin Cancer Res* 2010; 16: 1042–1048.
40. Ribas A, Hauschild A, Kefford R, et al. Phase III, open-label, randomized, comparative study of tremelimumab (CP-675,206) and chemotherapy (temozolomide [TMZ] or dacarbazine [DTIC]) in patients with advanced melanoma. *J Clin Oncol* 2008; 26(suppl 15S): abstract LBA9011.
41. Hamanishi J, Mandai M, Iwasaki M, et al. Programmed cell death 1 ligand 1 and tumor-infiltrating CD8+ T lymphocytes are prognostic factors of human ovarian cancer. *Proc Natl Acad Sci U S A* 2007; 104: 3360–3365.
42. Nishimura H, Nose M, Hiai H, et al. Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. *Immunity* 1999; 11: 141–151.
43. Nishimura H, Okazaki T, Tanaka Y, et al. Autoimmune dilated cardiomyopathy in PD-1 receptor-deficient mice. *Science* 2001; 291: 319–322.
44. Iwai Y, Ishida M, Tanaka Y, et al. Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade. *Proc Natl Acad Sci U S A* 2002; 99: 12293–12297.
45. Iwai Y, Terawaki S, Honjo T. PD-1 blockade inhibits hematogenous spread of poorly immunogenic tumor cells by enhanced recruitment of effector T cells. *Int Immunol* 2005; 17: 133–144.
46. Brahmer JR, Drake CG, Wollner I, et al. Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *J Clin Oncol* 2010; 28: 3167–3175.
47. Sznol M, Powderly JD, Smith DC, et al. Safety and antitumor activity of biweekly MDX-1106 (Anti-PD-1, BMS-936558/ONO-4538) in patients with advanced refractory malignancies. *J Clin Oncol* 2010; 28(suppl 15S): abstract 2506.
48. Brahmer JR, Topalian SL, Powderly J, et al. Phase II experience with MDX-1106 (Ono-4538), an anti-PD-1 monoclonal antibody, in patients with selected refractory or relapsed malignancies. *J Clin Oncol* 2009; 27(suppl 15S): abstract 3018.
49. US National Institutes of Health, ClinicalTrials.gov. Dose-escalation study of combination BMS-936558 (MDX-1106) and ipilimumab in subjects with unresectable stage III or stage IV malignant melanoma. http://www.clinicaltrials.gov/ct2/show/NCT01024231?term=ipilimumab&rank=29. Published December 2009. Accessed on July 20, 2010.
50. US National Institutes of Health, ClinicalTrials.gov. Vaccine therapy and monoclonal antibody therapy in treating patients with stage III or stage IV melanoma that cannot be removed by surgery. http://www.clinicaltrials.gov/ct2/show/NCT01176461?term=mdx+1106&rank=7. Published August 2010. Accessed on July 20, 2010.
51. Melero I, Hervas-Stubbs S, Glennie M, et al. Immunostimulatory monoclonal antibodies for cancer therapy. *Nat Rev Cancer* 2007; 7: 95–106.

52. Li B, Lin J, Vanroey M, et al. Established B16 tumors are rejected following treatment with GM-CSF-secreting tumor cell immunotherapy in combination with anti-4-1BB mAb. *Clin Immunol* 2007; 125: 76–87.
53. Choi BK, Kim YH, Kang WJ, et al. Mechanisms involved in synergistic anticancer immunity of anti-4-1BB and anti-CD4 therapy. *Cancer Res* 2007; 67: 8891–8899.
54. Sznol M, Hodi FS, Margolin K, et al. Phase I study of BMS-663513, a fully human anti-CD137 agonist monoclonal antibody, in patients with advanced cancer. *J Clin Oncol* 2008; 26(suppl): abstract 3007.
55. Khalil M, Vonderheide RH. Anti-CD40 agonist antibodies: preclinical and clinical experience. *Update Cancer Ther* 2007; 2: 61–65.
56. French RR, Chan HT, Tutt AL, et al. CD40 antibody evokes a cytotoxic T-cell response that eradicates lymphoma and bypasses T-cell help. *Nat Med* 1999; 5: 548–553.
57. Diehl L, den Boer AT, Schoenberger SP, et al. CD40 activation in vivo overcomes peptide-induced peripheral cytotoxic T-lymphocyte tolerance and augments anti-tumor vaccine efficacy. *Nat Med* 1999; 5: 774–779.
58. Sotomayor EM, Borrello I, Tubb E, et al. Conversion of tumor-specific CD4+ T-cell tolerance to T-cell priming through in vivo ligation of CD40. *Nat Med* 1999; 5: 780–787.
59. Grewal IS, Flavell RA. CD40 and CD154 in cell-mediated immunity. *Annu Rev Immunol* 1998; 16: 111–135.
60. Young LS, Eliopoulos AG, Gallagher NJ, et al. CD40 and epithelial cells: across the great divide. *Immunol Today* 1998; 19: 502–506.
61. Rosenberg SA, Dudley ME. Adoptive cell therapy for the treatment of patients with metastatic melanoma. *Curr Opin Immunol* 2009; 21: 233–240.
62. Rosenberg SA, Restifo NP, Yang JC, et al. Adoptive cell transfer: a clinical path to effective cancer immunotherapy. *Nat Rev Cancer* 2008; 8: 299–308.
63. Morton DL, Eilber FR, Holmes EC, et al. BCG immunotherapy of malignant melanoma: summary of a seven-year experience. *Ann Surg* 1974; 180: 635–643.
64. Soiffer R, Hodi FS, Haluska F, et al. Vaccination with irradiated, autologous melanoma cells engineered to secrete granulocyte-macrophage colony-stimulating factor by adenoviral-mediated gene transfer augments antitumor immunity in patients with metastatic melanoma. *J Clin Oncol* 2003; 21: 3343–3350.
65. Riker AI, Jove R, Daud AI. Immunotherapy as part of a multidisciplinary approach to melanoma treatment. *Front Biosci* 2006; 11: 1–14.
66. Ray S, Chhabra A, Mehrotra S, et al. Obstacles to and opportunities for more effective peptide-based therapeutic immunization in human melanoma. *Clin Dermatol* 2009; 27: 603–613.
67. Preynat-Seauve O, Contassot E, Schuler P, et al. Melanoma-infiltrating dendritic cells induce protective antitumor responses mediated by T cells. *Melanoma Res* 2007; 17: 169–176.
68. Jacobs JF, Aarntzen EH, Sibelt LA, et al. Vaccine-specific local T cell reactivity in immunotherapy-associated vitiligo in melanoma patients. *Cancer Immunol Immunother* 2009; 58: 145–151.
69. Sabel MS. Immunotherapy in the adjuvant setting: glass half-empty or half-full? Presented at: American Society for Clinical Oncology Annual Meeting; June 2010; Chicago, IL.
70. Stein JA, Adler A, Czernobilsky B, et al. Cutaneous response to autologous tumor cells in breast cancer and melanoma patients by active immunization. In: F. Spreafico and R. Arnon, eds. *Tumor Associated Antigens and Their Specific Immune Response.* London, England: Academic Press; 1979; 343–353.
71. Hollinshead A, Arlen M, Yonemoto R, et al. Pilot studies using melanoma tumor-associated antigens (TAA) in specific-active immunochemotherapy of malignant melanoma. *Cancer* 1982; 49: 1387–1404.
72. Hsueh EC, Morton DL. Antigen-based immunotherapy of melanoma: canvaxin therapeutic polyvalent cancer vaccine. *Semin Cancer Biol* 2003; 13: 401–407.
73. Morton DL, Mozzillo N, Thompson JF, et al. An international, randomized, phase III trial of Bacillus Calmette-Guérine(BCG) plus allogeneic melanoma vaccine (MCV) or placebo after complete resection of melanoma metastatic to regional or distant sites. *J Clin Oncol* 2007; 25(suppl 18S): abstract 8508.
74. Mitchell MS, Liggett PE, Green RL, et al. Sustained regression of a primary choroidal melanoma under the influence of a therapeutic melanoma vaccine. *J Clin Oncol* 1994; 12: 396–401.
75. Sondak VK, Liu PY, Tuthill RJ, et al. Adjuvant immunotherapy of resected, intermediate-thickness, node-negative melanoma with an allogeneic tumor vaccine: overall results of a randomized trial of the Southwest Oncology Group. *J Clin Oncol* 2002; 20: 2058–2066.
76. Pilla L, Valenti R, Marrari A, et al. Vaccination: role in metastatic melanoma. *Expert Rev Anticancer Ther* 2006; 6: 1305–1318.
77. Sosman JA, Unger JM, Liu PY, et al. Adjuvant immunotherapy of resected, intermediate-thickness, node-negative melanoma with an allogeneic tumor vaccine: impact of HLA class I antigen expression on outcome. *J Clin Oncol* 2002; 20: 2067–2075.
78. Selvan SR, Carbonell DJ, Fowler AW, et al. Establishment of stable cell lines for personalized melanoma cell vaccine. *Melanoma Res* 2010; 20: 280–292.
79. Thomas MC, Greten TF, Pardoll DM, et al. Enhanced tumor protection by granulocyte-macrophage colony-stimulating factor expression at the site of an allogeneic vaccine. *Hum Gene Ther* 1998; 9: 835–843.
80. Dranoff G, Jaffee E, Lazenby A, et al. Vaccination with irradiated tumor cells engineered to secrete murine granulocyte-macrophage colony-stimulating factor stimulates potent, specific, and long-lasting anti-tumor immunity. *Proc Natl Acad Sci U S A* 1993; 90: 3539–3543.
81. Jaffee EM, Lazenby A, Meurer J, et al. Use of murine models of cytokine-secreting tumor vaccines to study feasibility and toxicity issues critical to designing clinical trials. *J Immunother Emphasis Tumor Immunol* 1995; 18: 1–9.
82. Voskens CJ, Strome SE, Sewell DA. Synthetic peptide-based cancer vaccines: lessons learned and hurdles to overcome. *Curr Mol Med* 2009; 9: 683–693.
83. Baurain J, Stas M, Hammouch F, et al. Association of primary melanoma ulceration and clinical benefit of adjuvant vaccination with tumor-specific antigenic

peptides. *J Clin Oncol* 2009; 27(suppl 15S): abstract 3022.

84. Kruit WH, van Ojik HH, Brichard VG, et al. Phase 1/2 study of subcutaneous and intradermal immunization with a recombinant MAGE-3 protein in patients with detectable metastatic melanoma. *Int J Cancer* 2005; 117: 596–604.
85. Ruiz M, Kobayashi H, Lasarte JJ, et al. Identification and characterization of a T-helper peptide from carcinoembryonic antigen. *Clin Cancer Res* 2004; 10: 2860–2867.
86. Slingluff CL, Petroni GR, Smolkin ME, et al. Immunogenicity for CD8+ and CD4+ T cells of 2 formulations of an incomplete freund's adjuvant for multipeptide melanoma vaccines. *J Immunother* 2010; 33: 630–638.
87. Filipazzi P, Pilla L, Patuzzo R, et al. Adjuvant multipeptide vaccination in high-risk early melanoma patients. *J Clin Oncol* 2008; 26(suppl 15S): abstract 3014.
88. Slingluff CL, Petroni GR, Olson WC, et al. Effect of granulocyte-macrophage conolong stimulating factor(GM-CSF) administered with a multipeptide vaccine on circulating CD8+ and CD4+ T-cell responses: outcome of a multicenter randomized trial. *J Clin Oncol* 2009; 27(suppl 15S): abstract 3008.
89. Rosenberg SA, Yang JC, Schwartzentruber DJ, et al. Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. *Nat Med* 1998; 4: 321–327.
90. Schwartzentruber DJ, Lawson D, Richards J, et al. A phase III multi-institutional randomized study of immunization with the gp100:209-217(210M) peptide followed by high-dose IL-2 compared with high-dose Il-2 alone in patients with metastatic melanoma. *J Clin Oncol* 2009; 27(suppl 18S): abstract CRA9011.
91. O'Day S, Hodi FS, McDermott DF, et al. A phase III, randomized, double-blinded, multicenter study comparing monotherapy with ipilimumab or gp100 peptide vaccine and the combination in patients with previously treated, unresectable stage III or IV melanoma. *J Clin Oncol* 2010; 28(suppl 18S): abstract 4.
92. Hodi FS, O'Day S, McDermott DF, et al. Re-induction with ipilimumab, gp100 peptide vaccine, or a combination of both from a phase III, randomized, double-blind, multicenter study of previously treated patients with unresectable stage III or IV melanoma. *J Clin Oncol* 2010; 28(suppl 15S): abstract 8509.
93. Tosti G, di Pietro A, Ferrucci PF, et al. HSPPC-96 vaccine in metastatic melanoma patients: from the state of the art to a possible future. *Expert Rev Vaccines* 2009; 8: 1513–1526.
94. Di Pietro A, Tosti G, Ferrucci PF, et al. Heat shock protein peptide complex 96-based vaccines in melanoma: how far we are, how far we can get. *Hum Vaccin* 2009; 5: 727–737.
95. Testori A, Richards J, Whitman E, et al. Phase III comparison of vitespen, an autologous tumor-derived heat shock protein gp96 peptide complex vaccine, with physician's choice of treatment for stage IV melanoma: the C-100-21 Study Group [published correction appears in *J Clin Oncol* 2008;26:3819]. *J Clin Oncol* 2008; 26: 955–962.
96. Livingston PO, Wong GY, Adluri S, et al. Improved survival in stage III melanoma patients with GM2 antibodies: a randomized trial of adjuvant vaccination with GM2 ganglioside. *J Clin Oncol* 1994; 12: 1036–1044.
97. Kirkwood JM, Ibrahim JG, Sosman JA, et al. High-dose interferon alfa-2b significantly prolongs relapse-free and overall survival compared with the GM2-KLH/QS-21 vaccine in patients with resected stage IIB-III melanoma: results of intergroup trial E1694/S9512/C509801. *J Clin Oncol* 2001; 19: 2370–2380.
98. Livingston PO, Ritter G, Srivastava P, et al. Characterization of IgG and IgM antibodies induced in melanoma patients by immunization with purified GM2 ganglioside. *Cancer Res* 1989; 49(24 part 1): 7045–7050.
99. Eggermont AM, Suciu S, Rutkowski P, et al; EORTC Melanoma Group. Randomized phase III trial comparing postoperative adjuvant gaglioside GM2-KLH/QS-21 vaccination versus observation in stage II (T3-T4N0M0) melanoma: final results of study EORTC 18961. *J Clin Oncol* 2010; 28(suppl 15S): abstract 8505.
100. Abdel-Wahab Z, DeMatos P, Hester D, et al. Human dendritic cells, pulsed with either melanoma tumor cell lysates or the gp100 peptide(280-288), induce pairs of T-cell cultures with similar phenotype and lytic activity. *Cell Immunol* 1998; 186: 63–74.
101. Figdor CG, de Vries IJ, Lesterhuis WJ, et al. Dendritic cell immunotherapy: mapping the way. *Nat Med* 2004; 10: 475–480.
102. Nestle FO, Alijagic S, Gilliet M, et al. Vaccination of melanoma patients with peptide-or tumor lysate-pulsed dendritic cells. *Nat Med* 1998; 4: 328–332.
103. Banchereau J, Palucka AK, Dhodapkar M, et al. Immune and clinical responses in patients with metastatic melanoma to CD34(+) progenitor-derived dendritic cell vaccine. *Cancer Res* 2001; 61: 6451–6458.
104. Schadendorf D, Ugurel A, Schuler-Thurner B, et al. Dacarbazine (DTIC) versus vaccination with autologous peptide-pulsed dendritic cells (DC) in first-line treatment of patients with metastatic melanoma: a randomized phase III trial of the DC study group of the DeCOG. *Ann Oncol* 2006; 17: 563–570.
105. Palucka AK, Ueno H, Goldberg G, et al. Long-term survival and long-lived immune memory in patients with metastatic melanoma vaccinated with melanoma-antigen loaded dendritic cells. *J Clin Oncol* 2010; 28(suppl 15S): abstract e19018.
106. Markowicz S, Nowecki ZI, Rutkowski P, et al. Adjuvant vaccination with melanoma-antigen pulsed dendritic cells (DCs) in stage III melanoma patients. *J Clin Oncol* 2009; 27(suppl 15S): abstract 9039.
107. Wilgenhof S, Van Nuffel AM, Benteyn D, et al. Therapeutic vaccination with an autologous TriMix-Dendritic cell vaccine combined with sequential interferon alfa-2b in patients with advanced melanoma. *J Clin Oncol* 2009; 27(suppl 15S): abstract 9024.
108. Sivendran S, Pan M, Kaufman HL, et al. Herpes simplex virus oncolytic vaccine therapy in melanoma. *Expert Opin Biol Ther* 2010; 10: 1145–1153.
109. Hersey P, Coates AS, McCarthy WH, et al. Adjuvant immunotherapy of patients with high-risk melanoma

using vaccinia viral lysates of melanoma: results of a randomized trial. *J Clin Oncol* 2002; 20: 4181–4190.

110. Kaufman HL, DeRaffele G, Divito J, et al. A phase I trial of intralesional rV-Tricom vaccine in the treatment of malignant melanoma. *Hum Gene Ther* 2001; 12: 1459–1480.
111. Kaufman HL, Cheung K, Haskall Z, et al. Clinical protocol: intra-lesional rF-B7.1 versus rF-TRICOM vaccine in the treatment of metastatic cancer. *Hum Gene Ther* 2003; 14: 803–827.
112. Kaufman HL, Cohen S, Cheung K, et al. Local delivery of vaccinia virus expressing multiple costimulatory molecules for the treatment of established tumors. *Hum Gene Ther* 2006; 17: 239–244.
113. Wallack MK, Sivanandham M, Balch CM, et al. A phase III randomized, double-blind multiinstitutional trial of vaccinia melanoma oncolysate-active specific immunotherapy for patients with stage II melanoma. *Cancer* 1995; 75: 34–42.
114. Eager R, Nemunaitis J. GM-CSF gene-transduced tumor vaccines. *Mol Ther* 2005; 12: 18–27.
115. Senzer NN, Kaufman HL, Amatruda T, et al. Phase II clinical trial of a granulocyte-macrophage colony-stimulating factor-encoding, second-generation oncolytic herpesvirus in patients with unresectable metastatic melanoma. *J Clin Oncol* 2009; 27: 5763–5771.
116. Tang DC, DeVit M, Johnston, SA. Genetic immunization is a simple method for eliciting an immune response. *Nature* 1992; 356: 152–154.
117. Ulmer JB, Donnelly JJ, Parker SE, et al. Heterologous protection against influenza by injection of DNA encoding a viral protein. *Science* 1993; 259: 1745–1749.
118. Tagawa ST, Lee P, Snively J, et al. Phase I study of intranodal delivery of a plasmid DNA vaccine for patients with Stage IV melanoma. *Cancer* 2003; 98: 144–154.

Genetics of Prostate Cancer Risk

Mark M. Pomerantz, MD and Matthew L. Freedman, MD

Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA

OUTLINE

- Evidence for a Genetic Basis of Prostate Cancer
- Hereditary Prostate Cancer and Linkage Analysis
- Genome-Wide Association Studies
  - **Background**
  - **Chromosome 8q24: Insights Into Prostate Cancer Etiology**
- Use of Single Nucleotide Polymorphisms for Identifying Men at Risk for Prostate Cancer
- Genetic Risk Markers and Clinical Outcome
- Direction of Future Studies
- Conclusion

ABSTRACT

For decades, physicians and researchers have recognized that family history is a significant risk factor for prostate cancer. The identification of the genes responsible for inherited risk, however, proved difficult. With the sequencing of the human genome and the completion of the initial phases of the International HapMap Project, the tools are available to scan the entire genome and find genetic markers for disease. Since 2006, more than 30 inherited variants strongly associated with prostate cancer have been reported. As the inherited component of the disease is revealed, efforts are ongoing to translate genetic findings into the clinic. ***Mt Sinai J Med 77:643–654, 2010.*** **©** *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Matthew L. Freedman
Department of Medical Oncology
Dana-Farber Cancer Institute
Boston, MA
Email: freedman@broadinstitute.org

Key Words: family history, genetic markers, human genome, linkage analysis, prostate cancer.

Prostate cancer (CaP) is the leading cancer diagnosis, after cutaneous malignancies, among males in the United States,[1] and is the second most common cancer diagnosis among men worldwide.[2] There are significant geographic and ethnic disparities in the incidence of disease. These disparities are due to several factors, including differences in rates of CaP screening, disparate environmental exposures, and genetic variation. Yet in certain regions with relatively few cases, such as Japan and Eastern Europe, incidence is rising.[2] An estimated 1 in 6 American men will face a diagnosis of CaP, and many more are presumed to harbor the disease.[3,4] Although the majority of men with CaP will have an indolent form of the disease, CaP remains the second leading cause of cancer-related death in the United States, expected to lead to approximately 30,000 deaths in 2010.[1]

The remarkable prevalence of the disease, coupled with its potential to cause significant morbidity and mortality, presents unique challenges for researchers and clinicians. What are the underlying causes of CaP? How do we identify those at highest risk of CaP, especially those at risk for aggressive forms of the disease? The answers to these questions have profound implications on disease prevention, screening, and treatment. The strong heritable component of CaP indicates a central role for inherited variation in the development of disease. The following manuscript reviews the extensive efforts over the past 20 years to describe the genetic underpinnings of CaP. Particular attention is paid to recent breakthroughs in cancer genomics and the proliferation of data emerging from genome-wide association studies.

EVIDENCE FOR A GENETIC BASIS OF PROSTATE CANCER

The epidemiology of CaP supports the notion that the disease has a strong familial component. Family

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20222

© 2010 Mount Sinai School of Medicine

history has been examined extensively as a CaP risk factor in multiple studies,[5–20] and the familial contribution to CaP appears to be higher than for other common malignancies, such as breast and colorectal cancer.[7,21]

> *The epidemiology of prostate cancer supports the notion that the disease has a strong familial component. Family history has been examined extensively as a risk factor in multiple studies, and the familial contribution to prostate cancer appears to be higher than for other common malignancies, such as breast and colorectal cancer.*

A meta-analysis of 33 epidemiologic studies confirmed that subjects with a first-degree relative with CaP are at 2.53-fold lifetime risk of disease (95% confidence interval [CI]: 2.24–2.85).[22] Risk increased to approximately 5-fold for those with ≥2 affected first-degree family members. Risk appears to increase when a first-degree relative is diagnosed at an earlier age. In addition, risk of early-onset disease (before age 65) is influenced by family history. Studies have also consistently shown an increased risk for those with an affected brother, compared with those with an affected father: 3.37-fold versus 2.17-fold in the meta-analysis.[22]

These trends were reflected in a more recent analysis from the Health Professionals Follow-up Study, a study composed of a large population of men screened for prostate-specific antigen (PSA), including 3695 CaP cases. In this cohort, family history was associated with 2.25-fold risk of CaP before age 65 years and a 1.67-fold risk for onset after age 65. Those with a first-degree relative diagnosed before age 65 were at 2.16-fold risk of CaP, whereas those with a relative diagnosed after age 65 were at 1.95-fold risk.[23]

These studies suggest the importance of family history and identify the features associated with highest risk for CaP: number of affected relatives, age of relatives at diagnosis, and an affected brother compared with an affected father. However, these retrospective series are limited in their ability to distinguish definitively between environmental and genetic factors contributing to disease. The familial component of disease risk is not necessarily inherited, as family members share similar lifestyles and exposures. The stronger association among those with a brother with CaP may have a genetic etiology; it may indicate a recessive or X-linked genetic component to disease. Alternatively, a man with a brother diagnosed with CaP may be more likely to undergo prostate cancer screening. Or, environmental exposure at a young age, a factor more likely to be shared among brothers than between fathers and sons, is primarily responsible for risk.

> *The risk of prostate cancer (CaP) increases with the number of affected relatives, the youth of relatives at diagnosis, and occurrence of the disease in a brother versus a father.*

Studies analyzing disease patterns in twins have greater ability to isolate genetic from environmental factors in the development of disease. Twins are presumably exposed to similar environmental factors into adulthood, but monozygotic twins are genetically identical, whereas dizygotic (fraternal) twins share only 50% of their genomes. A Veterans Administration study of >16,000 male twins in the United States demonstrated a concordance for CaP diagnosis of 27.1% for monozygotic twins versus 7.1% for dizygotic twins.[24] Similarly, in 2000, an analysis of 44,788 pairs of twins listed in Swedish, Danish, and Finnish twin registries revealed a 21.1% concordance for monozygotic twins and a 6.4% concordance for dizygotic twins.[25] Using a model developed to determine the effects of heritable versus environmental factors, heritable factors were estimated to account for 42% of prostate cancer risk: greater than for the other 10 cancers analyzed, including breast and colorectal cancers.[25]

HEREDITARY PROSTATE CANCER AND LINKAGE ANALYSIS

Family history and twin studies strongly implicate heredity in disease susceptibility, and there has been interest in families in which CaP clusters most intensely. These families could provide insight into genes involved in disease risk, analogous to BRCA-1 and BRCA-2 in breast cancer and the APC gene in colon cancer. In an analysis of several hundred families affected by CaP, a definition for hereditary prostate cancer (HPC) emerged: (1) a cluster of ≥3 first-degree relatives with CaP, or (2) CaP in each of 3 generations, or (3) 2 relatives with CaP diagnosed before age 55.[26] Several groups developed

In an analysis of several hundred families affected by CaP, a definition for hereditary CaP emerged: (1) a cluster of ≥ 3 first-degree relatives with CaP, or (2) CaP in each of 3 generations, or (3) 2 relatives with CaP diagnosed before age 55.

genetic models based on Mendelian inheritance to account for familial aggregation of disease in CaP families (Figure 1). Using a method called segregation analysis, an early study included 691 families affected by prostate cancer and determined that hereditary predisposition could be due to a rare, autosomal dominant, highly penetrant allele.[27] The model accounted for 43% of early-onset CaP (before age 55) and 9% of all CaP diagnoses before age 85. Later series, examining thousands of CaP families, confirmed the possibility of this mode of inheritance.[28,29] Other segregation analyses favored an autosomal recessive or X-linked model.[30,31] More recently, MacInnis *et al.* used segregation analysis to determine the mode of inheritance in a United Kingdom and Australian population-based series of 4390 families, including 705 CaP cases.[32] The model accounting for a large proportion of genetic susceptibility in this cohort was composed of at least one recessively inherited allele conferring substantial risk, along with several more common genetic variants, each conferring small effects in a multiplicative fashion.

As HPC families were identified and modes of inheritance suggested, linkage studies were performed. Linkage studies screen the genome using genetic markers to identify regions that are cosegregating with the affected family members. The genetic regions linked within pedigrees plausibly contain a causal genetic variant.

Linkage mapping has produced several putative CaP risk loci. Among the most thoroughly analyzed is chromosome 1q24-25, originally reported in a 1996 study of 91 US and Swedish CaP families.[33] Other series have suggested that the locus, dubbed HPC1, is involved in disease affecting younger men with CaP (aged <66 years) with multiple affected family members.[34] Germline mutations in the tumor-suppressor gene RNASEL, located at HPC1, have been reported to track in CaP families, and have been implicated in up to 13% of all CaP cases.[35,36] The findings regarding HPC1 and RNASEL, however, have not been confirmed in independent series.[37–39] A meta-analysis of 10 studies examining RNASEL polymorphisms suggested no convincing influence on disease risk by these variants.[40] A locus at chromosome 22q has also been observed in several familial CaP series.[41–46] Johanneson *et al.* mapped the putative risk locus to a 15-kb region that includes the 5′ end of the gene *APOL3*.[46] Other risk loci have been implicated by linkage analysis,[47–52] but few have been consistently validated in independent cohorts.

Several factors complicate this type of study in CaP. Because CaP occurs relatively late in life, it is difficult to obtain DNA samples from >1 generation. Also, as CaP is a common disease, HPC families may include members who have developed sporadic forms of the disease. These subjects, termed "phenocopies," can confound linkage analysis. Finally, unlike BRCA-associated breast cancer,

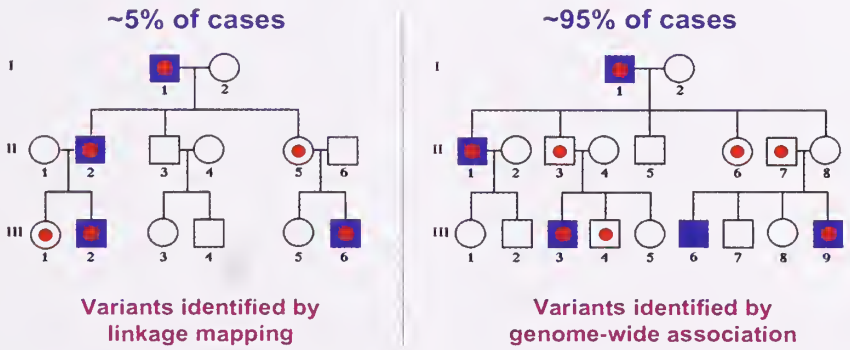

~5% of cases
~95% of cases
I
II
III
Variants identified by linkage mapping
Variants identified by genome-wide association

Fig 1. Two models for inherited prostate cancer risk. Three generations are depicted in each pedigree. Squares represent males and circles represent females. Blue squares depict those with disease, and those with a red dot carry a risk allele. The pedigree on the left demonstrates an autosomal dominant mode of inheritance. A risk allele of the type depicted in this model would be highly penetrant and carried by a small minority of patients. The pedigree on the right demonstrates the pattern of inheritance for common risk alleles, the type revealed by genome-wide association studies. These alleles are weakly penetrant.

which presents relatively early in life,[53] or familial adenomatous polyposis, which has a distinctive clinical presentation,[54] there is little to clinically or pathologically distinguish HPC from sporadic disease. Several series demonstrate no significant difference in clinicopathological features or outcome between familial and sporadic cases.[55–57]

GENOME-WIDE ASSOCIATION STUDIES

Background

Based on the paucity of validated findings in linkage analysis, one or even a handful of highly penetrant, rare alleles are unlikely to be responsible for a substantial proportion of CaP risk in the population. More likely, CaP is driven by multiple genetic and environmental contributions, a model consistent with the segregation analysis conducted by MacInnis *et al.*[32] Until recently, resources to identify common, less-penetrant genetic variants were limited. The completion of the Human Genome Project and the initial phase of the International HapMap Project made it possible to interrogate the entire genome for common markers of disease risk.[58] The approach, called a genome-wide association study (GWAS), is guided by the hypothesis that several common genetic variants, each with a modest effect, underlie common disease (Figure 1).

Most GWAS published to date scan the genome for single base variants, called single nucleotide polymorphisms (SNPs), that are associated with a trait of interest. More than 23 million SNPs have been catalogued in the public dbSNP database.[59,60] Genome-wide association studies compare a group of unrelated individuals with a condition of interest with a group of unrelated controls. The SNPs are interrogated individually to determine whether a variant allele in a particular SNP is disproportionately present in cases or in controls. "Common" SNPs, present in >1% of the human population, have been the focus of most GWAS to date. Interrogation of the entire catalog of common SNPs is unnecessary due to a fundamental property of the genome, termed linkage disequilibrium (LD). Nearby SNPs are often co-inherited more often than would be expected by chance. A single SNP can therefore serve as proxy for much of the variation in the surrounding genomic space, and the number of SNPs needed for a typical GWAS is greatly reduced. Genotyping 500,000 to 1 million "tagged" SNPs can capture roughly 80% of all common SNPs.[61]

Nonetheless, testing up to 1 million independent SNPs raises important statistical considerations.[62] Due to the potential for a large number of false positives, strict statistical thresholds are necessary to distinguish true positives from associations observed merely by chance. A P value threshold $<1 \times 10^{-7}$ is commonly applied. For common, polygenic diseases such as CaP, odds ratios (ORs) associated with risk alleles are generally within the range of 1.1 to 1.5. Large datasets, comprising thousands of cases and controls, are therefore necessary to identify variants conferring these modest effects.

Published GWAS have effectively identified >30 inherited variants strongly associated with CaP, primarily in men of European ancestry.[63–76] Table 1 lists CaP polymorphisms discovered to date reaching genome-wide statistical significance. These GWAS discoveries are the starting point for providing insight into the mechanisms of prostate carcinogenesis and identifying those at risk for disease.

Chromosome 8q24: Insights Into Prostate Cancer Etiology

In 2006, two independent GWAS of CaP, the first of their kind, were published. The two studies, using different methodologies, converged on the same chromosomal locus, 8q24.[63,64] Freedman *et al.* included 1597 African American CaP cases and 873 African American controls using a method called admixture mapping.[64] A 3.8-megabase (Mb) region of 8q24 (125.68–129.48) was identified as significantly associated with CaP risk. The region contains 9 known genes, including the well-described oncogene MYC. Investigators at deCODE Genetics also identified 8q24 as a risk locus in both European and African American populations.[63] Specific genetic polymorphisms highly associated with CaP risk in this study were situated

Genome-wide association studies (GWAS) have effectively identified >30 inherited variants strongly associated with CaP, primarily in men of European ancestry. In 2006, 2 independent GWAS of CaP, the first of their kind, were published. The 2 studies, using different methodologies, converged on the same chromosomal locus, 8q24.

Table 1. *Prostate Cancer Risk SNPs Reaching Statistical Significance in Genome-Wide Association Studies.*

SNP	Chromosomal locus	Nearest known gene within 100kb	Region
rs721048	2p15	EHBP1	intronic
rs1465618	2p21	THADA	intronic
rs12621278	2q31	ITGA6	intronic
rs2660753	3p12	VGLL3	intergenic
rs10934853	3q21	EEFSEC	intronic
rs12500426	4q22	PDLIM5	intronic
rs17021918	4q22	PDLIM5	intronic
rs7679673	4q24	TET2	intergenic
rs9364554	6q25	SLC22A3	intronic
rs10486567	7p15	JAZF1	intronic
rs6465657	7q21	LMTK2	intronic
rs2928679	8p21	SLC25A37	intergenic
rs1512268	8p21	NKX3-1	intergenic
rs10086908	8q24	none	intergenic
rs7841060	8q24	none	intergenic
rs13254738	8q24	none	intergenic
rs16901979	8q24	none	intergenic
rs16902094	8q24	none	intergenic
rs445114	8q24	none	intergenic
rs620861	8q24	none	intergenic
rs6983267	8q24	none	intergenic
rs7000448	8q24	none	intergenic
rs1447295	8q24	none	intergenic
rs10993994	10q11	MSMB	intergenic
rs4962416	10q26	CTBP2	intronic
rs7127900	11p15	TH	intergenic
rs11228565	11q13	MYEOV	intergenic
rs7931342	11q13	MYEOV	intergenic
rs11649743	17q12	HNF1B	intronic
rs4430796	17q12	HNF1B	intronic
rs1859962	17q24	none	intergenic
rs8102476	19q13	PPP1R14A	intergenic
rs2735839	19q13	KLK3	intergenic
rs5759167	22q13	BIK	intergenic
rs5945619	Xp11	NUDT11	intergenic

Table 2. *Prostate Cancer Risk SNPs Reaching Statistical Significance in Genome-Wide Association Studies (part 2).*

SNP	Risk allele frequency[a]	Odds ratio[b]	Reference
rs721048	0.198–0.240	1.15	71
rs1465618	0.180–0.210	1.16–1.20	74
rs12621278	0.940	1.32–1.47	74
rs2660753	0.118–0.128	1.11–1.48	69
rs10934853	0.291–0.330	1.12	75
rs12500426	0.560	1.14–1.17	74
rs17021918	0.630–0.650	1.12–1.25	74
rs7679673	0.610–0.620	1.15–1.37	74
rs9364554	0.323 -0.328	1.17–1.26	69
rs10486567	0.770	1.12–1.35	70
rs6465657	0.492–0.508	1.03–1.19	69
rs2928679	0.450–0.460	1.16–1.26	74
rs1512268	0.400–0.420	1.13–1.28	74
rs10086908	0.720–0.740	1.14–1.25	73
rs7841060	0.246	1.19	72
rs13254738	0.348	1.11	67
rs16901979	0.034–0.073	1.66	65
rs16902094	0.127–0.248	1.21	75
rs445114	0.613–0.710	1.14	75
rs620861	0.610–0.662	1.11–1.28	72, 73
rs6983267	0.530–0.580	1.13–1.42	67, 68, 70, 73
rs7000448	0.400	1.14	67
rs1447295	0.074–0.169	1.29–1.72	63, 65, 68
rs10993994	0.458–0.459	1.15–1.42	69
rs4962416	0.270	1.17–1.20	70
rs7127900	0.180	1.29–1.40	74
rs11228565	0.210–0.291	1.23	75
rs7931342	0.556–0.562	1.19–1.25	69
rs11649743	0.800–0.860	1.28	76
rs4430796	0.469–0.580	1.16–1.38	66, 76
rs1859962	0.489–0.522	1.20	66
rs8102476	0.473–0.641	1.12	75
rs2735839	0.871	1.25–1.72	69
rs5759167	0.500–0.530	1.14–1.20	74
rs5945619	0.398–0442	1.19–1.46	69

within the 3.8 Mb region defined by Freedman *et al.* Several groups subsequently validated these findings in geographically and ethnically diverse cohorts[7–79]

The prostate cancer–related causal alleles in 8q24 have been defined more precisely in several mapping studies. The studies have defined 5 distinct risk loci, involving at least 9 single nucleotide polymorphisms, all independently associated with prostate cancer risk.

Risk variants discovered by GWAS are merely associated with a trait of interest and are not necessarily the causal polymorphisms. Risk SNPs are likely in LD with causal variants, accounting for their identification in GWAS. To isolate the causal alleles at 8q24 more precisely, Haiman *et al.* initiated a fine mapping study of the 3.8-Mb region to close in on the disease-causing alleles at 8q24.[67] The study defined 3 distinct risk regions: region 1, spanning 8q24 128.54–128.62 Mb; region 2, 128.14–128.28 Mb; and region 3, 128.470–128.54 Mb. In 2009, Al Olama *et al.* and Yeager *et al.* discovered 2 additional risk regions at 8q24.[72,73] At least 9 SNPs, all independently associated with CaP risk, reside within these 5 loci (Figure 2). Among carriers of the risk alleles, ORs for disease are estimated to range from 1.1 to 2.0 in populations of European ancestry.[67,72,73] Chromosome 8q24 is particularly intriguing because, in addition to CaP, risk markers for breast, colon, and bladder cancer and chronic lymphocytic leukemia have been discovered there.[80–84]

Notably, all 8q24 risk polymorphisms reside in intergenic, noncoding regions of the genome. Therefore, the mechanism by which inherited risk variation leads to cancer is unknown. A leading hypothesis is that the risk alleles influence regulation of a distal gene. The nearest genes to 8q24 risk loci are MYC and FAM84B, each more than 250 kilobases (kb) from the nearest risk SNP (Figure 2). Consistent with this hypothesis, the risk loci appear to bear the epigenetic marks of genetic enhancers, elements capable of affecting expression of 1 or more genes from long range.[85] However, RNA expression of the genes at 8q24, including MYC, does not appear correlated with risk allele status in prostate tissue.[86] Because it is a known oncogene, MYC nonetheless remains an attractive candidate for involvement in CaP risk. With this in mind, a technique to detect long-range chromatin looping, called chromosome conformation capture (3C), was performed from risk regions 1 to 3. Interactions between the risk loci and 36 segments, spanning >1 Mb across 8q24, were measured. Highly significant interactions between each of the risk loci and MYC, across hundreds of kb, were clearly identified[87] (Figure 2). These findings suggest involvement by MYC in CaP risk. Outstanding questions include how and when regulation of MYC takes place, but a mechanism for 8q24-associated risk is beginning to emerge. The findings may provide a paradigm for other risk regions, as the majority of CaP risk loci are intergenic. The discoveries at

> *Notably, all 8q24 risk polymorphisms reside in intergenic, noncoding regions of the genome. Therefore, the mechanism by which inherited risk variation leads to cancer is unknown.*

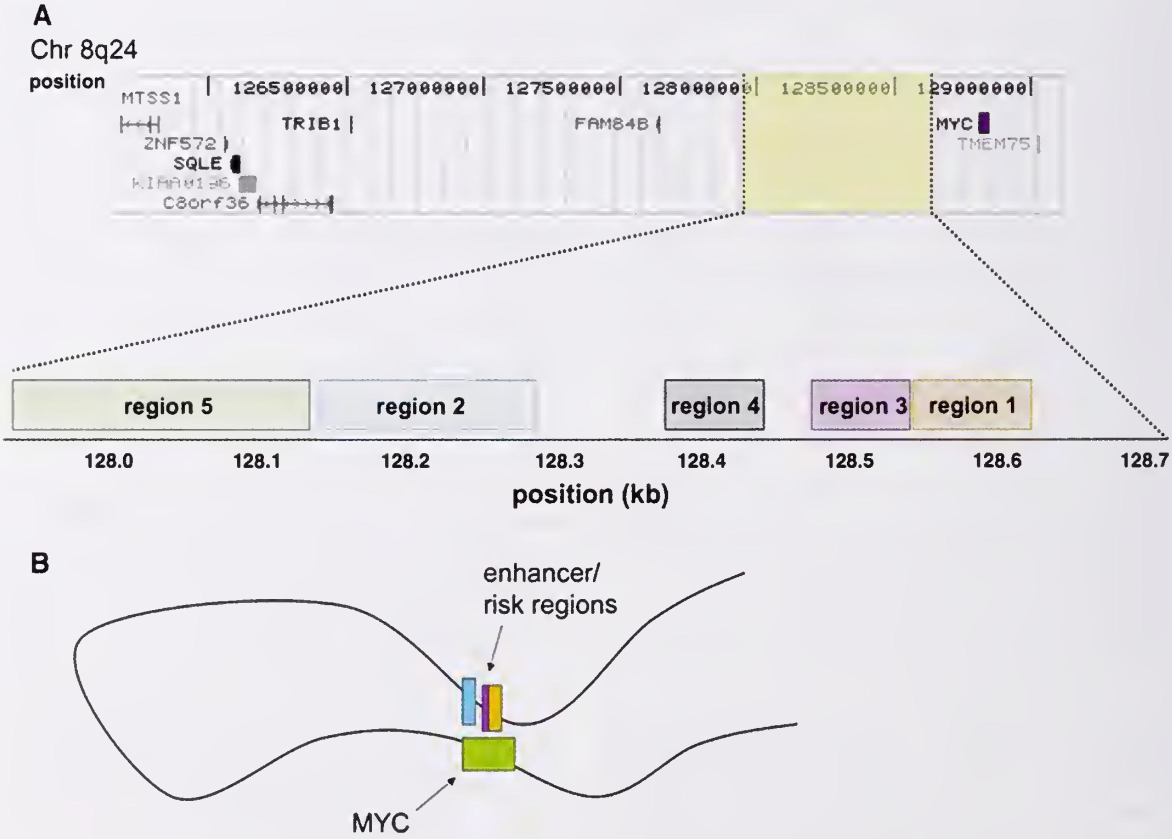


Fig 2. The chromosome 8q24 prostate cancer risk locus. (A) The image at the top depicts a 3-Mb region at 8q24, including the genes in the area and the gene desert, in yellow, where the risk variants reside. Beneath are displayed the 5 prostate cancer risk loci within the gene desert. (B) A rough depiction of the looping revealed by the 3C assay. Data from 3C assays suggest that the risk loci, which contain genetic enhancers, are in close proximity to MYC in 3-dimensional space despite being hundreds of kb away. **Abbreviations:** 3C, chromosome conformation capture.

8q24 demonstrate the potential for GWAS results to elucidate the underpinnings of inherited risk.

USE OF SINGLE NUCLEOTIDE POLYMORPHISMS FOR IDENTIFYING MEN AT RISK FOR PROSTATE CANCER

Results from GWAS could prove beneficial to CaP patients as predictors of disease risk. Germline genetic markers have a natural advantage over current screening tools, such as PSA, because they are static: they are ever-present and do not fluctuate with time or clinical condition. Whereas PSA only reaches clinical attention when CaP has presumably already developed, inherited risk SNPs are testable at any time prior to the presence of disease. These considerations must be balanced, however, against how much information is actually gained by their inclusion. Some studies have been performed in this area; however, more work is necessary before a clear consensus emerges.

Zheng *et al.* examined the clinical utility of genetic profiling of 2893 men in their mid-60s with CaP and 1781 age-matched controls for 5 independent risk variants in the first set of CaP risk loci discovered: 8q24, 17q12, and 17q24.[88] Risk of CaP correlated with increasing number of risk alleles inherited. When family history was added to the profile as a sixth risk marker, the correlation grew significantly stronger. For men carrying 5 of the possible 6 risk factors (5 SNPs plus family history), the OR was 9.46 for developing CaP when compared with those with no risk allele and no family history of disease. However, when the authors constructed receiver operating characteristic curves to measure the sensitivity and specificity of genomic profiling, the area under the curve (AUC) was 0.63 for a profile involving age, geographic region, family history, and genotype of the 5 risk SNPs. This was a very modest improvement over an AUC of 0.61 when genotype was eliminated. This same group added SNPs to their model as they were discovered by GWAS. When 11 risk SNPs were later analyzed, AUC improved to only 0.65,[89] still well below 0.8, the threshold generally considered to represent accurate prediction.[90] More recently, in a population of 1809 Icelandic cases and 783 controls, 22 known CaP variants were genotyped. The 1.3% of men considered at highest risk were 2.5-fold more likely to develop disease by age 75 compared with the average Icelandic man.[75]

A major determinant limiting the potential of genetic profiling for CaP is its polygenic nature. The contribution from each risk variant is modest (OR: 1.22–1.53 in Zheng *et al.*), and, as a result, a large majority of subjects will be at average risk of disease. In Zheng *et al.*, 85.5% of cases and 86.2% of controls genotyped harbored 1–3 risk markers.[88] As a result, finding a proper "cutoff" for declaring high or low risk becomes impossible without accepting unreasonable numbers of false positives or false negatives.

> *A major determinant limiting the potential of genetic profiling for prostate cancer is its polygenic nature. Zheng* et al. *found the contribution from each risk variant to be modest (OR: 1.22–1.53). In addition, 85.5% of cases and 86.2% of controls genotyped harbored 1–3 risk markers. Thus, finding a proper "cutoff" for declaring high or low risk becomes impossible without accepting unreasonable numbers of false positives or false negatives.*

It has been proposed, then, that genetic profiling based on GWAS findings should be aimed at finding those at highest risk of disease. Yet, even for this small subset, <5% of men, clinical benefit of profiling is not clear. Those considered high risk, for example, could be recommended for earlier PSA screening, as is often advised for those with a family history of disease. There are no known modifiable risk factors for CaP that could be recommended to high-risk patients, though chemoprevention strategies, such finasteride, may be reasonable. Finasteride, a 5-alpha-reductase inhibitor, has been demonstrated to prevent or delay the onset of CaP in men aged >55 years.[91] Future clinical trials could target carriers of multiple risk alleles and determine whether this population particularly benefits from chemoprevention.

GENETIC RISK MARKERS AND CLINICAL OUTCOME

If genetic profiling were able to predict clinical course, clinicians would have a useful tool to help guide treatment decisions. The majority of men with CaP in the United States have localized disease at the time of diagnosis. However, not all men require treatment or exposure to the morbidity associated with radical prostatectomy or radiation therapy. Germline genetic markers that accurately distinguish aggressive from nonaggressive disease could have a significant impact on patient care. There does appear to be a genetic component to CaP outcome,[92] and several groups have examined whether the CaP risk SNPs discovered to date predict outcome.

As risk SNPs have been discovered via GWAS, several groups have genotyped large cohorts for these variants to determine whether they are associated with clinical variables. Because clinical annotation is limited for most large retrospective series, the variables available for analysis in such studies are most often surrogates for outcome: age at diagnosis, PSA at diagnosis, Gleason score, stage, and some combination of these variables.

Only a handful of studies have been able to directly evaluate associations between risk allele status and CaP-specific mortality, arguably the most critical and clinically useful variable. Penney *et al.* and Salinas *et al.* analyzed the first set of risk markers reported by GWAS, at chromosome 8q24 and 17q, and found no correlation between genotype and survival.[93,94] Wiklund *et al.* analyzed a cohort that included 440 CaP-related deaths, genotyping a set of 16 risk SNPs, and did not detect an association with mortality.[95] Most recently, Gallagher *et al.* interrogated 29 risk SNPs in a cohort of 798 cases that included 91 CaP-specific deaths. The group detected a significant association between the SNP rs2735839 and CaP-specific mortality ($P = 0.0007$).[96] Interestingly, the *nonrisk* allele was associated with increased risk of death from CaP. In a larger study that included >4000 CaP cases and 883 deaths, the nonrisk allele at rs2735839 was also significantly associated CaP-specific mortality (Pomerantz *et al.*, unpublished data). Consistent with this finding, other series have demonstrated that the protective rs2735839 allele is associated with higher Gleason score and increased disease aggressiveness.[97,98]

> *Only a handful of studies have been able to directly evaluate associations between risk allele status and prostate cancer–specific mortality, arguably the most critical and clinically useful variable. Though 2 studies have found no correlation, Gallagher* et al. *detected a significant association between prostate cancer–specific mortality and the rs2735839 ($P = 0.0007$).*

The risk SNP rs2735839 resides in the intergenic region of chromosome 19q13, approximately 600 base pairs downstream of the 3′UTR of KLK3 and approximately 12.2 kilobases upstream of the transcription start site of KLK2. KLK3 (also known as kallikrein 3) encodes PSA, and polymorphisms at 19q13, including rs2735839, have been associated with PSA levels,[69,98–103] specifically in men without a diagnosis of CaP. When rs2735839 was reported as a risk SNP for CaP, it was speculated that the finding was a result of ascertainment bias due to its association with PSA. In this scenario, those inheriting the risk allele would have a slightly higher PSA and would be more likely to receive a referral for biopsy. As more men with this genotype are biopsied, more men with this genotype are diagnosed with CaP. Likewise, it is possible that ascertainment bias is involved in the association of the nonrisk allele with aggressiveness. In a series comparing indolent with aggressive CaP cases, the indolent group is more likely to be enriched with individuals who were biopsied based on a slightly elevated PSA.

Validation in a non–CaP-screened cohort may help resolve the issue, because the data would not be confounded by screening PSAs. It is possible that there are two variants, one associated with risk of CaP and the other associated with lethal disease, each linked to rs2735839 and having differential downstream effects. If the variant is associated with outcome, prospective series could also determine whether genotype at this variant can help guide treatment decisions. If rs2735839 is associated solely with PSA level, consideration could be made regarding PSA threshold for biopsy referral. Carriers of the risk allele, whose PSA trends higher, would have a higher cutoff for "normal," compared with carriers of the nonrisk allele, who would have a lower cutoff.

DIRECTION OF FUTURE STUDIES

While GWAS still hold great promise, data generated to date are clearly limited in their clinical applicability. The current collection of known, common CaP variants accounts for only a fraction of the genetic component of CaP risk. Perhaps more variants will be found in future GWAS and in meta-analyses of past studies that will help make genetic profiling more informative. It is presumed that rare variants, polymorphisms with minor allele frequencies <1%, account for a substantial portion, maybe a majority, of inherited risk. As these variants are discovered through efforts such as the 1000 Genomes Project and studied in large GWAS, they may have significantly higher ORs for disease compared with common SNPs and may enable more effective profiling. In addition, there are other types of genetic polymorphisms, such as copy number variants, that are becoming increasingly amenable to testing. Broadening genome-wide searches beyond just SNPs may yield useful markers.

Also, it is critical that GWAS are performed across multiple ethnic groups. Most GWAS conducted to date involve subjects of European ancestry. For a disease such as CaP that varies in incidence in different ancestral populations, this limits the number of potentially important variants. For example, risk allele frequency within certain loci at 8q24 is significantly higher in African Americans, and one African American risk SNP, Broad11934905 at region 2, is not found in European American, Latino, Japanese American, or Hawaiian populations. This SNP appears to confer a higher relative risk of disease than any other marker at 8q24.[67]

Finally, further work is needed to discover genetic polymorphisms associated with outcome. Disease aggressiveness has a heritable component,[92] but inherited variants associated with risk of disease are not necessarily the alleles responsible for aggressive CaP. A GWAS dedicated to aggressive and nonaggressive CaP could address this important issue. Further GWAS focused on response to treatments such as hormonal therapy or chemotherapy could also guide clinical decisions.

CONCLUSION

Since the initial observations that CaP runs in families, the heritability of CaP has definitively been established. Over the past 4 years, >30 bona fide germline variants have been discovered that partially explain the heritable component of the disease. The next immediate steps include deciphering the remainder of the genetic factors responsible for CaP and finding ways to translate these discoveries into the clinic.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Jemal A, Siegel R, Xu J, et al. Cancer statistics, 2010. *CA Cancer J Clin* 2010; 60: 277–300.
2. Jemal A, Center MM, DeSantis C, et al. Global patterns of cancer incidence and mortality rates and trends. *Cancer Epidemiol Biomarkers Prev* 2010; 19: 1893–1907.
3. Sakr WA, Grignon DJ, Haas GP, et al. Age and racial distribution of prostatic intraepithelial neoplasia. *Eur Urol* 1996; 30: 138–144.
4. Taichman RS, Loberg RD, Mehra R, et al. The evolving biology and treatment of prostate cancer. *J Clin Invest* 2007; 117: 2351–2361.
5. Steinberg GD, Carter BS, Beaty TH, et al. Family history and the risk of prostate cancer. *Prostate* 1990; 17: 337–347.
6. Spitz MR, Currier RD, Fueger JJ, et al. Familial patterns of prostate cancer: a case-control analysis. *J Urol* 1991; 146: 1305–1307.
7. Goldgar DE, Easton DF, Cannon-Albright LA, et al. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst* 1994; 86: 1600–1608.
8. Aprikian AG, Bazinet M, Plante M, et al. Family history and the risk of prostatic carcinoma in a high risk group of urological patients. *J Urol* 1995; 154(2 part 1): 404–406.
9. Hayes RB, Liff JM, Pottern LM, et al. Prostate cancer risk in U.S. blacks and whites with a family history of cancer. *Int J Cancer* 1995; 60: 361–364.
10. Isaacs SD, Kiemeney LA, Baffoe-Bonnie A, et al. Risk of cancer in relatives of prostate cancer probands. *J Natl Cancer Inst* 1995; 87: 991–996.
11. Narod SA, Dupont A, Cusan L, et al. The impact of family history on early detection of prostate cancer. *Nat Med* 1995; 1: 99–101.
12. Grönberg H, Damber L, Damber JE. Familial prostate cancer in Sweden: a nationwide register cohort study. *Cancer* 1996; 77: 138–143.
13. Lesko SM, Rosenberg L, Shapiro S. Family history and prostate cancer risk. *Am J Epidemiol* 1996; 144: 1041–1047.
14. Ghadirian P, Howe GR, Hislop TG, et al. Family history of prostate cancer: a multi-center case-control study in Canada. *Int J Cancer* 1997; 70: 679–681.
15. Rodríguez C, Calle EE, Miracle-McMahill HL, et al. Family history and risk of fatal prostate cancer. *Epidemiology* 1997; 8: 653–657.
16. Schuurman AG, Zeegers MP, Goldbohm RA, et al. A case-cohort study on prostate cancer risk in relation to family history of prostate cancer. *Epidemiology* 1999; 10: 192–195.
17. Cerhan JR, Parker AS, Putnam SD, et al. Family history and prostate cancer risk in a population-based cohort of Iowa men. *Cancer Epidemiol Biomarkers Prev* 1999; 8: 53–60.
18. Hemminki K, Czene K. Age specific and attributable risks of familial prostate carcinoma from the family-cancer database. *Cancer* 2002; 95: 1346–1353.
19. Valeri A, Cormier L, Moineau MP, et al. Targeted screening for prostate cancer in high risk families: early onset is a significant risk factor for disease in first degree relatives. *J Urol* 2002; 168: 483–487.
20. Bruner DW, Moore D, Parlanti A, et al. Relative risk of prostate cancer for men with affected relatives:

systematic review and meta-analysis. *Int J Cancer* 2003; 107: 797–803.
21. Hemminki K, Sundquist J, Bermejo JL. How common is familial cancer? *Ann Oncol* 2008; 19: 163–167.
22. Zeegers MP, Jellema A, Ostrer H. Empiric risk of prostate carcinoma for relatives of patients with prostate carcinoma: a meta-analysis. *Cancer* 2003; 97: 1894–1903.
23. Chen YC, Page JH, Chen R, et al. Family history of prostate and breast cancer and the risk of prostate cancer in the PSA era. *Prostate* 2008; 68: 1582–1591.
24. Page WF, Braun MM, Partin AW, et al. Heredity and prostate cancer: a study of World War II veteran twins. *Prostate* 1997; 33: 240–245.
25. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer-analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000; 343: 78–85.
26. Carter BS, Bova GS, Beaty TH, et al. Hereditary prostate cancer: epidemiologic and clinical features. *J Urol* 1993; 150: 797–802.
27. Carter BS, Beaty TH, Steinberg GD, et al. Mendelian inheritance of familial prostate cancer. *Proc Natl Acad Sci U S A* 1992; 89: 3367–3371.
28. Grönberg H, Damber L, Damber JE, et al. Segregation analysis of prostate cancer in Sweden: support for dominant inheritance. *Am J Epidemiol* 1997; 146: 552–557.
29. Schaid DJ, McDonnell SK, Blute ML, et al. Evidence for autosomal dominant inheritance of prostate cancer. *Am J Hum Genet* 1998; 62: 1425–1438.
30. Pakkanen S, Baffoe-Bonnie AB, Matikainen MP, et al. Segregation analysis of 1,546 prostate cancer families in Finland shows recessive inheritance. *Hum Genet* 2007; 121: 257–267.
31. Monroe KR, Yu MC, Kolonel LN, et al. Evidence of an X-linked or recessive genetic component to prostate cancer risk. *Nat Med* 1995; 1: 827–829.
32. MacInnis RJ, Antoniou AC, Eeles RA, et al. Prostate cancer segregation analyses using 4390 families from UK and Australian population-based studies. *Genet Epidemiol* 2010; 34: 42–50.
33. Smith JR, Freije D, Carpten JD, et al. Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. *Science* 1996; 274: 1371–1374.
34. Grönberg H, Xu J, Smith JR, et al. Early age at diagnosis in families providing evidence of linkage to the hereditary prostate cancer locus (HPC1) on chromosome 1. *Cancer Res* 1997; 57: 4707–4709.
35. Carpten J, Nupponen N, Isaacs S, et al. Germline mutations in the ribonuclease L gene in families showing linkage with HPC1. *Nat Genet* 2002; 30: 181–184.
36. Casey G, Neville PJ, Plummer SJ, et al. RNASEL Arg462Gln variant is implicated in up to 13% of prostate cancer cases. *Nat Genet* 2002; 32: 581–583.
37. Eeles RA, Durocher F, Edwards S, et al. The Cancer Research Campaign/British Prostate Group U.K. Familial Prostate Cancer Study Collaborators. Linkage analysis of chromosome 1q markers in 136 prostate cancer families. *Am J Hum Genet* 1998; 62: 653–658.
38. Berry R, Schaid DJ, Smith JR, et al. Linkage analyses at the chromosome 1 loci 1q24-25 (HPC1), 1q42.2-43 (PCAP), and 1p36 (CAPB) in families with hereditary prostate cancer. *Am J Hum Genet* 2000; 66: 539–546.
39. Xu J. Combined analysis of hereditary prostate cancer linkage to 1q24-25: results from 772 hereditary prostate cancer families from the International Consortium for Prostate Cancer Genetics [published correction appears in *Am J Hum Genet* 2000; 67: 541–542. *Am J Hum Genet* 2000; 66: 945–957.
40. Li H, Tai BC. RNASEL gene polymorphisms and the risk of prostate cancer: a meta-analysis. *Clin Cancer Res* 2006; 12: 5713–5719.
41. Cunningham JM, McDonnell SK, Marks A, et al. Genome linkage screen for prostate cancer susceptibility loci: results from the Mayo Clinic Familial Prostate Cancer Study. *Prostate* 2003; 57: 335–346.
42. Janer M, Friedrichsen DM, Stanford JL, et al. Genomic scan of 254 hereditary prostate cancer families. *Prostate* 2003; 57: 309–319.
43. Chang BL, Isaacs SD, Wiley KE, et al. Genome-wide screen for prostate cancer susceptibility genes in men with clinically significant disease. *Prostate* 2005; 64: 356–361.
44. Stanford JL, McDonnell SK, Friedrichsen DM, et al. Prostate cancer and genetic susceptibility: a genome scan incorporating disease aggressiveness. *Prostate* 2006; 66: 317–325.
45. Camp NJ, Farnham JM, Cannon-Albright LA. Localization of a prostate cancer predisposition gene to an 880-kb region on chromosome 22q12.3 in Utah high-risk pedigrees. *Cancer Res* 2006; 66: 10205–10212.
46. Johanneson B, McDonnell SK, Karyadi DM, et al. Family-based association analysis of 42 hereditary prostate cancer families identifies the Apolipoprotein L3 region on chromosome 22q12 as a risk locus. *Hum Mol Genet* 2010; 19: 3852–3862.
47. Berthon P, Valeri A, Cohen-Akenine A, et al. Predisposing gene for early-onset prostate cancer, localized on chromosome 1q42.2-43. *Am J Hum Genet* 1998; 62: 1416–1424.
48. Xu J, Meyers D, Freije D, et al. Evidence for a prostate cancer susceptibility locus on the X chromosome. *Nat Genet* 1998; 20: 175–179.
49. Gibbs M, Stanford JL, McIndoe RA, et al. Evidence for a rare prostate cancer-susceptibility locus at chromosome 1p36. *Am J Hum Genet* 1999; 64: 776–787.
50. Stanford JL, FitzGerald LM, McDonnell SK, et al. Dense genome-wide SNP linkage scan in 301 hereditary prostate cancer families identifies multiple regions with suggestive evidence for linkage. *Hum Mol Genet* 2009; 18: 1839–1848.
51. Berry R, Schroeder JJ, French AJ, et al. Evidence for a prostate cancer-susceptibility locus on chromosome 20. *Am J Hum Genet* 2000; 67: 82–91.
52. Lange EM, Beebe-Dimmer JL, Ray AM, et al. Genome-wide linkage scan for prostate cancer susceptibility from the University of Michigan Prostate Cancer Genetics Project: suggestive evidence for linkage at 16q23. *Prostate* 2009; 69: 385–391.
53. Welcsh PL, King MC. BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. *Hum Mol Genet* 2001; 10: 705–713.
54. Rustgi AK. The genetics of hereditary colon cancer. *Genes Dev* 2007; 21: 2525–2538.

55. Bova GS, Partin AW, Isaacs SD, et al. Biological aggressiveness of hereditary prostate cancer: long-term evaluation following radical prostatectomy. *J Urol* 1998; 160: 660–663.
56. Grönberg H, Damber L, Tavelin B, et al. No difference in survival between sporadic, familial and hereditary prostate cancer. *Br J Urol* 1998; 82: 564–567.
57. Roehl KA, Loeb S, Antenor JA, et al. Characteristics of patients with familial versus sporadic prostate cancer. *J Urol* 2006; 176: 2438–2442.
58. The International HapMap Consortium. The International HapMap Project. *Nature* 2003; 426: 789–796.
59. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001; 29: 308–311.
60. Sayers EW, Barrett T, Benson DA, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 2010; 38: D5–D16.
61. Kruglyak L. The road to genome-wide association studies. *Nat Rev Genet* 2008; 9: 314–318.
62. Wang WY, Barratt BJ, Clayton DG, et al. Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 2005; 6: 109–118.
63. Amundadottir LT, Sulem P, Gudmundsson J, et al. A common variant associated with prostate cancer in European and African populations. *Nat Genet* 2006; 38: 652–658.
64. Freedman ML, Haiman CA, Patterson N, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. *Proc Natl Acad Sci U S A* 2006; 103: 14068–14073.
65. Gudmundsson J, Sulem P, Manolescu A, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet* 2007; 39: 631–637.
66. Gudmundsson J, Sulem P, Steinthorsdottir V, et al. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat Genet* 2007; 39: 977–983.
67. Haiman CA, Patterson N, Freedman ML, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. *Nat Genet* 2007; 39: 638–644.
68. Yeager M, Orr N, Hayes RB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 2007; 39: 645–649.
69. Eeles RA, Kote-Jarai Z, Giles GG, et al. Multiple newly identified loci associated with prostate cancer susceptibility. *Nat Genet* 2008; 40: 316–321.
70. Thomas G, Jacobs KB, Yeager M, et al. Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet* 2008; 40: 310–315.
71. Gudmundsson J, Sulem P, Rafnar T, et al. Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. *Nat Genet* 2008; 40: 281–283.
72. Yeager M, Chatterjee N, Ciampa J, et al. Identification of a new prostate cancer susceptibility locus on chromosome 8q24. *Nat Genet* 2009; 41: 1055–1057.
73. Al Olama AA, Kote-Jarai Z, Giles GG, et al. Multiple loci on 8q24 associated with prostate cancer susceptibility. *Nat Genet* 2009; 41: 1058–1060.
74. Eeles RA, Kote-Jarai Z, Al Olama AA, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat Genet* 2009; 41: 1116–1121.
75. Gudmundsson J, Sulem P, Gudbjartsson DF, et al. Genome-wide association and replication studies identify four variants associated with prostate cancer susceptibility. *Nat Genet* 2009; 41: 1122–1126.
76. Sun J, Zheng SL, Wiklund F, et al. Evidence for two independent prostate cancer risk-associated loci in the HNF1B gene at 17q12. *Nat Genet* 2008; 40: 1153–1155.
77. Schumacher FR, Feigelson HS, Cox DG, et al. A common 8q24 variant in prostate and breast cancer from a large nested case-control study. *Cancer Res* 2007; 67: 2951–2956.
78. Severi G, Hayes VM, Padilla EJ, et al. The common variant rs1447295 on chromosome 8q24 and prostate cancer risk: results from an Australian population-based case-control study. *Cancer Epidemiol Biomarkers Prev* 2007; 16: 610–612.
79. Suuriniemi M, Agalliu I, Schaid DJ, et al. Confirmation of a positive association between prostate cancer risk and a locus at chromosome 8q24. *Cancer Epidemiol Biomarkers Prev* 2007; 16: 809–814.
80. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007; 447: 1087–1093.
81. Tomlinson I, Webb E, Carvajal-Carmona L, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet* 2007; 39: 984–988.
82. Zanke BW, Greenwood CM, Rangrej J, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* 2007; 39: 989–994.
83. Kiemeney LA, Thorlacius S, Sulem P, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. *Nat Genet* 2008; 40: 1307–1312.
84. Crowther-Swanepoel D, Broderick P, Di Bernardo MC, et al. Common variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 influence chronic lymphocytic leukemia risk. *Nat Genet* 2010; 42: 132–136.
85. Jia L, Landan G, Pomerantz M, et al. Functional enhancers at the gene-poor 8q24 cancer-linked locus. *PLoS Genet* 2009; 5: e1000597.
86. Pomerantz MM, Beckwith CA, Regan MM, et al. Evaluation of the 8q24 prostate cancer risk locus and MYC expression. *Cancer Res* 2009; 69: 5568–5574.
87. Ahmadiyeh N, Pomerantz MM, Grisanzio C, et al. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. *Proc Natl Acad Sci U S A* 2010; 107: 9742–9746.
88. Zheng SL, Sun J, Wiklund F, et al. Cumulative association of five genetic variants with prostate cancer. *N Engl J Med* 2008; 358: 910–919.
89. Zheng SL, Sun J, Wiklund F, et al. Genetic variants and family history predict prostate cancer similar to prostate-specific antigen. *Clin Cancer Res* 2009; 15: 1105–1111.
90. Janssens AC, Gwinn M, Bradley LA, et al. A critical appraisal of the scientific basis of commercial genomic profiles used to assess health risks and personalize health interventions. *Am J Hum Genet* 2008; 82: 593–599.
91. Thompson IM, Goodman PJ, Tangen CM, et al. The influence of finasteride on the development of prostate cancer. *N Engl J Med* 2003; 349: 215–224.

92. Hemminki K, Ji J, Forsti A, et al. Concordance of survival in family members with prostate cancer. *J Clin Oncol* 2008; 26: 1705–1709.
93. Penney KL, Salinas CA, Pomerantz M, et al. Evaluation of 8q24 and 17q risk loci and prostate cancer mortality. *Clin Cancer Res* 2009; 15: 3223–3230.
94. Salinas CA, Koopmeiners JS, Kwon EM, et al. Clinical utility of five genetic variants for predicting prostate cancer risk and mortality. *Prostate* 2009; 69: 363–372.
95. Wiklund FE, Adami HO, Zheng SL, et al. Established prostate cancer susceptibility variants are not associated with disease outcome. *Cancer Epidemiol Biomarkers Prev* 2009; 18: 1659–1662.
96. Gallagher DJ, Vijai J, Cronin AM, et al. Susceptibility loci associated with prostate cancer progression and mortality. *Clin Cancer Res* 2010; 16: 2819–2832.
97. Kader AK, Sun J, Isaacs SD, et al. Individual and cumulative effect of prostate cancer risk-associated variants on clinicopathologic variables in 5,895 prostate cancer patients. *Prostate* 2009; 69: 1195–1205.
98. Xu J, Isaacs SD, Sun J, et al. Association of prostate cancer risk variants with clinicopathologic characteristics of the disease. *Clin Cancer Res* 2008; 14: 5819–5824.
99. Ahn J, Berndt SI, Wacholder S, et al. Variation in KLK genes, prostate-specific antigen and risk of prostate cancer. *Nat Genet* 2008; 40: 1032–1036.
100. Cramer SD, Chang BL, Rao A, et al. Association between genetic polymorphisms in the prostate-specific antigen gene promoter and serum prostate-specific antigen levels. *J Natl Cancer Inst* 2003; 95: 1044–1053.
101. Lai J, Kedda MA, Hinze K, et al. PSA/KLK3 AREI promoter polymorphism alters androgen receptor binding and is associated with prostate cancer susceptibility. *Carcinogenesis* 2007; 28: 1032–1039.
102. Severi G, Hayes VM, Neufing P, et al. Variants in the prostate-specific antigen (PSA) gene and prostate cancer risk, survival, and circulating PSA. *Cancer Epidemiol Biomarkers Prev* 2006; 15: 1142–1147.
103. Pal P, Xi H, Sun G, et al. Tagging SNPs in the kallikrein genes 3 and 2 on 19q13 and their associations with prostate cancer in men of European origin. *Hum Genet* 2007; 122: 251–259.

Malignant Gliomas: New Translational Therapies

Joohee Sul, MD and Howard A. Fine, MD

Neuro-Oncology Branch, National Cancer Institute, National Institutes of Neurological Disorders
and Stroke/National Institutes of Health, Bethesda, MD

OUTLINE

Classification of Malignant Gliomas
Epigenetics
Targeted Molecular Therapies
  Epidermal Growth Factor Pathway
  Antiangiogenesis and Vascular Endothelial Growth Factor
Immunotherapy
  Dendritic Cell Vaccines
  Peptide Vaccines
Stem Cell Therapy
Conclusion

ABSTRACT

Malignant gliomas are the most common primary brain tumors in adults and carry a dismal prognosis. Despite aggressive therapy with maximal safe surgical resection, radiation and chemotherapy, these tumors invariably are refractory to or become resistant to treatment and recur. Gliomas are highly infiltrative cancers and display remarkable genetic heterogeneity making them challenging to treat. Recent progress has been made in understanding the molecular and genetic composition of these tumors and from this, promising new targets for therapy have emerged. In particular, anti-angiogenesis therapies have led to modest success in disease control. In addition, the growing body of research in cancer immunology as well as cancer stem cells has made inroads in our understanding of tumorgenesis. Translational research has been particularly crucial to the development of these therapies as much preclinical and clinical work is needed to develop the rationale for treatments, to develop biomarkers of drug activity and to elucidate mechanisms of resistance. This brief overview will discuss some of the pivotal advances made in the pursuit of improved outcomes and survival for patients with this devastating disease. ***Mt Sinai J Med 77:655–666, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: angiogenesis, epigenetics, glioblastoma, glioma stem cells, immunotherapy, malignant glioma, molecular profiles, signal transduction pathways, translational research, tyrosine kinase inhibitors.

Malignant gliomas are the most common primary brain tumors in adults, with an incidence of 7 per 100,000 in the United States.[1] The World Health Organization (WHO) grading system is a widely accepted scale used to categorize these tumors, with the 2 chief histopathologies being anaplastic astrocytoma (AA), WHO grade III tumors; and glioblastoma (GBM), WHO grade IV.[2] Anaplastic oligodendrogliomas and mixed oligoastrocytomas are also WHO grade III tumors, but less common than AAs. Glioblastomas are the most aggressive of the primary brain tumors and exhibit steadfast resistance to treatment. Although survival has improved for many other solid tumors, it remains poor for patients with malignant gliomas. The current standard of care for GBM was determined by a large phase III trial published in 2005 conducted by the European Organization for Research and Treatment of Cancer (EORTC) and National Cancer Institute of Canada (NCIC).[3] Stupp *et al.* randomized patients with newly diagnosed glioblastomas to either radiation therapy (RT) alone or RT with temozolomide, an oral chemotherapy. Patients who received combination therapy had improved progression-free survival (PFS) and overall survival (OS) when compared with the RT-alone group (Table 1). More recent follow-up from this trial corroborates the durability of these findings. In patients receiving both RT and chemotherapy, 3-, 4-, and 5-year survival was improved compared with those receiving RT alone

Address Correspondence to:

Howard A. Fine
Neuro-Oncology Branch,
NCI/NIH
Bethesda, MD
Email: hfine@mail.nih.gov

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20223

© 2010 Mount Sinai School of Medicine

Table 1. *Outcomes by Randomized Treatment Regimen for Newly Diagnosed Glioblastomas*[3].

	Treatment (mo)	
Outcomes	RT	RT + TMZ
OS	12.1	14.6
PFS	5.0	6.9

Abbreviations: OS, overall survival; PFS, progression-free survival; RT, radiation therapy; TMZ, temozolomide.

Table 2. *Long-Term Follow-Up of the EORTC-NCIC Trial*[4].

	Survival (%)			
Treatment	2-year	3-year	4-year	5-year
RT	10.9	4.4	3.0	1.9
RT + TMZ	27.2	16.0	12.1	9.8

Abbreviations: EORTC-NCIC, European Organisation for Research and Treatment of Cancer-National Cancer Institute of Canada; RT, radiation therapy; TMZ, temozolomide.

(Table 2).[4] Although the addition of chemotherapy proved beneficial in these patients, progression of disease remains nearly universal.

Several factors contribute to the resistance of these tumors to treatment. Gliomas are diffuse by nature, with infiltrating tumor cells invading adjacent areas of normal brain. Thus, surgical resection is inadequate for local control, and these reservoirs of residual tumor undoubtedly lead to recurrent disease. Although somewhat sensitive to high-dose radiation, treatment of gliomas with radiotherapy is limited by normal tissue toxicity to the normal brain. Chemotherapy, the linchpin of treatment for most cancers, has had only modest effects on improving outcomes for gliomas. This is in part due to the protective nature of the blood brain barrier (BBB), which can preferentially exclude chemotherapeutics and other targeted therapies. In addition, malignant gliomas are among the most genetically heterogeneous cancers, making it difficult to achieve durable responses.

Fortunately, significant advances in basic and clinical research have been made in recent years toward understanding the genetics of these tumors, as well as the abnormal cell-signaling pathways that lead to these cancers. This brief review will highlight some of the recent developments that have resulted from the interchange between preclinical and clinical efforts. Substantial research has been devoted to elucidating the genetic and molecular profiles of gliomas, and new therapies, such as those directed against angiogenesis, have appreciably changed the outcomes and course of this disease. Basic research in tumor immunology and cancer stem cells has also contributed to our understanding of gliomagenesis. Yet despite promising results in the preclinical setting, many of these treatments have not yet resulted in the anticipated clinical responses. As such, translational research is particularly valuable to understanding why the discoveries that occur in the laboratory have not yet been borne out in the clinic.

CLASSIFICATION OF MALIGNANT GLIOMAS

Brain tumors have classically been diagnosed and graded by pathological evaluation of cellular elements. Anaplastic astrocytomas are characterized by pleomorphic cells, nuclear atypia, and mitosis. Glioblastomas exhibit these same features, as well as necrosis and vascular proliferation. Yet despite having consistent histologic descriptions, patient outcomes and response to treatment can be quite varied. It is clear that the current classification systems are inadequate to fully characterize these tumors. Underlying genetic variability exists among these tumors, and it is likely that each has a distinctive profile. It is hoped that identifying subgroups among malignant gliomas will lead to more predictable responses to treatment within these groups, and to more personalized therapies for individual patients.

> *Current classification systems for malignant gliomas are inadequate to fully characterize them. Underlying genetic variability exists among these tumors, and it is likely that each has a distinctive profile. Identifying subgroups among malignant gliomas will hopefully lead to more predictable responses to treatment within these groups, and to more personalized therapies for individual patients.*

There have been large-scale efforts made toward understanding the genetic and molecular profiles of gliomas, and glioblastomas in particular. The Cancer Genome Atlas is a coordinated effort of the National Cancer Institut[3] and the National Human Genome Research Institute aimed at determining the principal genome alterations leading to malignancy. Glioblastomas were the first cancers to be studied under The Cancer Genome Atlas.[5] Interpretation of DNA, mRNA, and micro RNA profiles of hundreds of primary glioblastoma tissue samples is ongoing and has led to the categorization of these tumors into 4 main types: classical, mesenchymal, neural, and proneural.[6] Validation of these subgroups is challenging, and the clinical applications are as yet uncertain.

The Glioma Molecular Diagnostic Initiative is another large-scale effort to develop a classification schema based on unbiased gene-expression profiles. Previous classification schemas have been stratified by a priori histological classes or candidate genes. In contrast, the Glioma Molecular Diagnostic Initiative relies only on molecular data without premeditated stratification to create its framework. From this, 2 major groups of gliomas are predicted: glioblastoma-rich and oligodendroglioma-rich. These have been further separated into 6 subtypes.[7] These data are being collected to be assimilated into a comprehensive molecular/clinical database for public use.[8] The efforts toward unlocking the molecular signatures of malignant gliomas are in progress with the ultimate goal of defining a new classification model with both prognostic and predictive value.

EPIGENETICS

Epigenetics refers to changes in gene expression or phenotype that do not occur as a result of changes in DNA. Methylation of DNA is a primary mechanism of epigenetic changes in gene expression, and it has been recognized that hypomethylation of genes occurs in many malignancies. The DNA repair protein O6-methylguanine-DNA methyltransferase (MGMT) has been implicated in resistance to chemotherapeutic agents,[9–11] and epigenetic silencing of the MGMT gene by methylation of its promoter region has been correlated with better responses to alkylating agents.[12,13] Methylation leads to loss of MGMT expression and chemotherapy-induced damage to malignant cells is left unrepaired, leading to apoptosis. An important corollary to the previously described EORTC study reports MGMT methylation occurring in 46 of 206 patients.[13] This was found to be an independent prognostic factor, and OS was significantly improved in those patients with promoter methylation who received combination therapy compared with those patients who received RT alone or who were not methylated (see Table 3).

Table 3. *O6-Methylguanine-DNA Methyltransferase Promoter Methylation Status and OS for Newly Diagnosed Glioblastomas*[13].

	OS (mo)	
Promoter Status	RT	RT + TMZ
Methylated	15.3	21.7
Unmethylated	11.8	12.7

Abbreviations: OS, overall survival; RT, radiation therapy; TMZ, temozolomide.

Several strategies have been employed in an attempt to overcome MGMT expression, including adding O6-benzylguanine (O6-BG), an irreversible inhibitor of MGMT, to chemotherapy. Although preclinical studies were promising,[14,15] clinical trials using this agent have not shown clear benefit.[16–18] Quinn *et al.* reported results from a phase II study of temozolomide in combination with O6-BG in patients with recurrent malignant gliomas on temozolomide. The overall objective response rate was 9%, and the median PFS was 7.9 weeks in this group.[16] The disappointing clinical outcomes may in part be due to the additive toxicity of O6-BG to cytotoxic agents, potentiating bone marrow suppression and requiring substantial dose reductions. In this study, the dose of temozolomide was decreased by 50%, and in another phase II study using carmustine in combination of O6-BG, an 80% dose reduction in carmustine was necessary.[18]

Another strategy to overcome MGMT-mediated chemotherapy resistance is the use of alternative dosing schedules. Dose-dense or metronomic schedules both lead to more prolonged exposure to temozolomide, and can potentially deplete MGMT. Tolcher *et al.* used 2 different dosing schedules (7 days on/7 days off or 21/28 days consecutively), and took serial blood samples to evaluate for MGMT activity.[19] Both dosing schedules resulted in reducing MGMT activity, but with recovery of function during the rest period. Clarke *et al.* conducted a randomized phase II trial of dose-dense or metronomic temozolomide following standard chemoradiation in newly diagnosed glioblastomas.[20] There was some suggestion that patients with unmethylated MGMT may have benefited from the dose-dense regimen because their median OS was improved compared with those unmethylated patients in the EORTC trial (15.4 months versus 12.7 months).

Methylation status of MGMT has not yet been fully validated as a molecular marker of response, and as such current recommendations for temozolomide treatment are not based on this parameter. A large phase III trial conducted by the Radiation Therapy Oncology Group is currently under way comparing standard temozolomide dosing schedule to 21/28-day dosing, and patients have been stratified by MGMT methylation status prior to randomization. The results of this multi-institutional trial may help to validate the clinical significance of MGMT status.

TARGETED MOLECULAR THERAPIES

Intracellular signal transduction pathways determine all the key aspects of tumor biology, including proliferation, motility, apoptosis, and angiogenesis.

> *Intracellular signal transduction pathways determine all the key aspects of tumor biology, including proliferation, motility, apoptosis, and angiogenesis.*

Many of these pathways are regulated by the interaction of growth factors and their cell-surface receptors. Tyrosine kinases (TKs) are a major class of membrane receptors that are often deregulated in cancer, and therefore they have become prominent targets of pathway inhibition. Most agents that target TKs are either small molecule TK inhibitors (TKIs) or monoclonal antibodies (mAbs). As the growth pathways for malignant gliomas are identified, corresponding targeted therapies have been developed. Two major pathways have emerged as promising candidates: the epidermal growth factor (EGF) and its receptor, EGFR; and the angiogenesis pathway involving vascular endothelial growth factor (VEGF). Vascular endothelial growth factor is a key mediator of the angiogenesis pathway and EGFR is a promoter of multiple factors in cell proliferation, invasiveness, and angiogenesis.

Epidermal Growth Factor Pathway

Overexpression or overactivity of EGFR has been associated with numerous malignancies and has become a prominent focus in cancer research. Activation of EGFR launches a host of downstream signaling cascades leading to cell proliferation, motility, and adhesion. In malignant gliomas, EGFR signaling is increased via overexpression or mutation in 40%–50% of all tumors, but not so in normal brain tissue.[21,22] In addition, EGFRvIII is a specific EGFR mutant that is constitutively active, is present in 20%–30% of glioblastomas, and may be associated with worse OS.[23–25] Some preclinical data suggest that EGFR inhibition slows tumor proliferation and inhibits glioma invasion. This profile makes EGFR an attractive target for glioma treatment. The most common approaches to anti-EGFR therapy include the use of mAbs, small-molecule TKIs, and peptide vaccination.

Small-molecule TKIs act by blocking adenosine triphosphate binding in the intracellular catalytic domain of EGFR tyrosine kinase. Agents such as gefitinib and erlotinib are prototypes of these drugs and have been studied in multiple clinical trials as single-agent therapy as well as in combination with standard chemotherapy for malignant gliomas (see Table 4). Rich *et al.* reported minimal responses with gefitinib, and no improvement in OS in 53 patients with recurrent glioblastomas.[26] In addition, EGFR expression on tumor tissue was not found to be a predictor of response or OS. Erlotinib has activity against both EGFR and EGFRvIII. Results with erlotinib have also been mixed, with studies reporting response rates of 0% to 25%.[27]

Table 4. *Phase II Clinical Trials With EGFR Inhibitors in Patients With Glioblastoma.*

Authors	No. of Patients	Radiographic Response Rates (%)	PFS (mo)	OS (mo)
Gefitinib				
Rich *et al.*[26]	57	0	2.0	9.9
Lieberman *et al.*[89]	38	13	2.0	—
Uhm *et al.*[90]	96	—	6.8	12.8
Franceschi *et al.*[91]	16	0	2.1	6.2
Erlotinib				
Vogelbaum *et al.*[92]	16	25	5.2	—
Raizer *et al.*[93]	31	0	2.3	—
Cloughesy *et al.*[94]	48	8	—	—
Van den Bent *et al.*[95]	54	4	1.8	7.7

Abbreviations: EGFR, epidermal growth factor receptor; OS, overall survival; PFS, progression-free survival.

These results demonstrate that the promise of EGFR inhibition in the preclinical setting has yet to translate to significant clinical responses. The reason for this is unclear, although curiously the current data suggest there is no clear correlation between EGFR expression and tumor response in the clinical setting.[26] One potential explanation for this observation is that there are other genetic determinants that modify the susceptibility of glioma cells to EGFR inhibition. For instance, there is some evidence that tumors with EGFRvIII and intact phosphatase and tensin homolog (PTEN; a tumor-suppresser gene) are more susceptible to EGFR inhibitors; however, this has yet to be validated.[28] As is true of all agents, another confounding variable in the clinical effectiveness of EGFR inhibitors is drug delivery. Both the anatomical microarchitecture of the blood-brain barrier (BBB) as well as drug efflux transporters on the endothelial cells that constitute the BBB combine to impede drug penetration into normal cerebral as well as tumor tissue. The extent of pathway inhibition is critical to determining a clinical response, and is dependent on the ability of the agent to reach its target in vivo.[29] Accurate measurements of drug levels in tumor tissue are difficult to assess in this population given the invasiveness of brain-tissue sampling. Finally, the activation of alternate pathways that bypass the EGFR signaling by direct downstream signaling activation likely also contributes to resistance to EGFR-targeted therapies. In fact, preclinical models have demonstrated that multiple kinase inhibition may be required to optimally suppress tumor growth.[30]

Antiangiogenesis and Vascular Endothelial Growth Factor

Contrary to the variable results we have seen with EGFR inhibition, the strategy of antiangiogenesis for malignant gliomas has yielded moderate success. Cancer-related blood vessels have been studied since the 1800s, but it was Dr. Judah Folkman's work in the early 1970s that established tumors were restricted in growth and would only reach a size of 1–2 mm^3 unless they were able to generate an autonomous blood supply.[31–33] Folkman recognized that inhibiting the growth of blood vessels could lead to effective cancer treatment and postulated that antiangiogenesis would be a viable avenue to tumor inhibition.

Malignant gliomas are highly vascularized, and as such are well suited for treatment with antiangiogenesis agents. The process is driven by multiple factors, including VEGF, platelet-derived growth factor (PDGF), stem cell factor-1, and basic fibroblast growth factor (bFGF).[34–36] The VEGF pathway, however, is clearly central to the promotion of neovascularization and it has been the principal target of most antiangiogenic therapies.

The mechanism of action for anti-VEGF therapy is likely multifold and has not been clearly proven. Two of the more commonly discussed hypotheses include (1) direct inhibition of endothelial cell proliferation, leading to decreased blood vessel formation and hypoxic cell death, and (2) normalization of the tumor vasculature, leading to decreased interstitial pressure and subsequent enhanced delivery of chemotherapeutic agents.[35,37] Batchelor *et al.* presented early clinical and imaging data supporting the use of antiangiogenesis agents in glioblastoma.[38] Using AZD2171, a TKI against the VEGF receptor, they demonstrated normalization of tumor blood vessels in patients with recurrent GBMs. The radiographic response rate was 56%, and patients were often able to reduce or discontinue steroid use as a result of dramatic reduction in peritumoral edema. The authors also measured blood vessel size, permeability, contrast enhancement, and edema at serial time points using dynamic contrast-enhanced magnetic resonance imaging (MRI). These factors were correlated with serum biomarkers of response such as circulating endothelial and progenitor cells. As a corollary to these findings, a recent follow-up to this same study describes changes in several serum biomarkers: VEGF, placental growth factor, and stromal cell–derived factor-1α were increased after VEGF blockade, whereas matrix metalloproteinase (MMP)-10 was decreased. An increase in both plasma MMP-2 and urinary MMP-9/neutrophil gelatinase–associated lipocalin was associated with poorer outcomes.[39]

Bevacizumab is a humanized mAb against VEGF and was the first commercially available angiogenesis inhibitor. It was approved by the US Food and Drug Administration in February 2004 for use in metastatic colorectal cancer when used with standard chemotherapy treatment. Early data on the activity of bevacizumab in malignant gliomas was reported in 2005.[40] An unprecedented 9 of 21 patients with recurrent malignant glioma had radiographic response with the combination of bevacizumab and CPT-11, a topoisomerase inhibitor. Following this, there have been multiple studies demonstrating the utility of bevacizumab in treating malignant gliomas (see Table 5).[40–43] In 2009, bevacizumab received accelerated approval from the FDA for treatment of recurrent glioblastomas, based on 2 phase II trials.[41,42] The use of VEGF inhibitors has significantly impacted outcomes for patients with progressive disease or patients who have failed standard therapy.

Table 5. *Clinical Trials With Bevacizumab in Patients With Malignant Gliomas.*

Authors	No. of patients	Radiographic Response Rates (%)	PFS (mo)	OS (mo)
Stark-Vance[40]	21	43	—	—
Friedman *et al.*[41]	Bev alone: 85	28	4.2	9.2
	Bev + CPT-11: 82	38	5.6	8.7
Kreisl *et al.*[42]	48	35	4.0	—
Vredenburgh *et al.*[43]	35	57	—	—

Abbreviations: Bev, bevacizumab; CPT, irinotecan; OS, overall survival; PFS, progression-free survival.

In the wake of the responses seen with bevacizumab, other inhibitors of angiogenesis have been studied, including TKIs targeted against VEGF, PDGF, and stem cell factor-1.

The use of vascular endothelial growth factor inhibitors has significantly impacted outcomes for patients with progressive disease or patients who have failed standard therapy.

Despite the relative success in improving response rates with VEGF inhibitors, gliomas ultimately escape the effects of antiangiogenesis therapies. Insight into the mechanisms of resistance has shown that levels of several growth factors, including bFGF and stromal cell derived factor-1α were increased at the time of tumor progression.[38] These pathways may serve as alternative routes for angiogenesis in the presence of VEGF inhibition. There are also emerging data that blood vessel stability relies on pericyte-endothelial cell interactions that are promoted by PDGF. In the setting of anti-VEGF therapy PDGF signaling is increased, and thereby is a potential target for therapy in conjunction with VEGF inhibitors.[44] Finally, there is increasing evidence to suggest that anti-VEGF therapy alone may trigger a distinctive pattern of recurrence characterized by enhanced infiltrative growth and co-opting of existing blood vessels.[45–48] Thus, it is highly plausible that even if complete antiangiogenic blockade can be achieved, tumor progression will eventually occur through individual glioma cell invasion along preexisting normal cerebral vasculature. Thus, antiangiogenic treatment may change the pathogenesis of end-stage disease from one of a large tumor mass causing mass effect, cerebral edema, and focal cerebral destruction to a diffuse encephalic clinical picture as a result of widely infiltrative individual glioma cells, as is seen in the condition gliomatosis cerebri.

IMMUNOTHERAPY

In theory, the immune system is the ideal mechanism for the treatment of highly infiltrative gliomas. Ideally, the host's innate immune cells would identify specific tumor antigens and launch a directed response against any disseminated cells while leaving normal cells unharmed. Strategies to exploit the immune system include passive immunotherapy using mAbs, or active immunotherapy using tumor vaccines. Both tactics depend on the successful activation of lymphocytes against tumor antigens. Although the intellectual rationale for these approaches is valid, significant clinical success has yet to be seen. Clinical trials using vaccines have been far more prevalent and are the main focus of this discussion.

The central nervous system (CNS) had long been thought of as an "immune-privileged" sanctuary separated from the general circulation by the BBB, and devoid of native immune cells or lymphoid tissue. It is now known that although the immune responses generated in the CNS are generally less efficient than systemic responses, there is ample activity, as witnessed by the vigorous immune reactions seen in the setting of CNS infections and autoimmune diseases. Although resident antigen-presenting cells (APCs) have yet to be identified it has been postulated that microglia may assume this role in the brain.[49,50] Early preclinical studies demonstrated that glioma cells expressing interferon-gamma transplanted into brain tissue lead to priming of microglia and tumor rejection.[51]

The model of immune activation is dependent on uptake of antigens by APCs, which then travel through lymphatic channels to reach lymph nodes. Once there, APCs with their associated antigen trigger activation of T cells. These primed lymphocytes then travel to the tumor and initiate cell death. The identification of many tumor-associated antigens has spurred the development of vaccine therapies mediated by T cells. There are numerous sources for vaccine antigens, including cancer cell lysates and specific peptide or proteins related to the cancer.

These antigens need not be limited to those involved with the oncogenesis of the tumor, but any antigen specific to malignant cells and absent in normal tissue. The 2 major types of vaccines discussed below, dendritic cell (DC) vaccines and peptide vaccines, have both been studied in clinical trials.

Dendritic Cell Vaccines

Dendritic cells function as APCs, and once activated they travel to lymphoid tissue where they initiate T cell and B cell responses. They are so named for their dendrite-like projections and have no relation to neurons. Most current DC vaccine strategy involves harvesting peripheral autologous DCs from patients and exposing them ex vivo to tumor cell antigens or lysates. These DCs are then reintroduced to the patients with the intent of inducing an endogenous immune response. Preclinical studies using DC vaccines have yielded encouraging results.[52,53] An early phase I trial proved the safety of dendritic cell vaccines in humans.[54] Patient-derived DCs and glioma cells were pulsed together, then injected intradermally. There were no clear tumor responses, but interferon-gamma expression was detected in the peripheral blood monocyte cells of a few patients. In another phase I dose-escalation study, Liau *et al.* treated 12 GBM patients with DCs pulsed with autologous tumor peptides.[55] Although the study was not powered to measure efficacy, patients did have improved PFS (15.5 months) and OS (23.4 months) compared with historical controls.

Although DC vaccine trials have gained popularity, there has been little proof of clinical benefit in the form of measurable MRI changes. It is uncertain if the responses seen in many of these trials may be restricted to specific a population. In the previous phase I study by Liau *et al.*, tumor burden and evidence of disease progression at the time of vaccination were significant factors in determining OS and presence of tumor infiltration by T cells.[55] Only those patients who had no radiographic evidence of disease progression at the time of vaccination were shown to mount a systemic cytotoxic T-cell response, and only those with minimal tumor burden exhibited tumor-infiltrating lymphocytes. Moreover, the actual ability of DCs to successfully prime T cells against tumor antigens is still in question. There is no direct evidence of the mechanism by which the T-cell priming may occur, and there has been some recent data that implicates DCs in the brain as inhibitors of T-cell priming.[56] Even if reactive T cells are primed against the proper antigens, they may have no impact on tumor growth. Rosenberg *et al.* demonstrated this in melanoma patients where normal tumor growth persisted despite the presence of up to 30% of T cells displaying antitumor activity.[57] The use of supplemental treatments to enhance the effects of DC vaccination has been explored, and includes the addition of cytokines[58] and adjuvant chemotherapy.[59]

Peptide Vaccines

A concern with the use of nonspecific autologous tumor extracts is the danger of inducing autoimmune disease.[60] To circumvent this risk, tumor-specific antigens have been explored as candidate targets. The advantage of peptide vaccines lies in the ease of access to synthetically produced tumor associated peptides, and the decreased risk of induced autoimmune states. The obvious disadvantage with this technique is the lack of specificity for individual gliomas, which can be widely heterogeneous. EGFRvIII is the prototype for a tumor-specific antigen, and an early phase I study with a vaccine using autologous DCs pulsed with the peptide fragment of EGFRvIII (PEPvIII) demonstrated improved OS in 12 newly diagnosed GBM patients.[61] Following this, preclinical data demonstrated that direct vaccination with the PEPvIII was sufficient to inhibit tumor growth, and subsequent phase II trials were conducted without DCs, using peptide-based vaccines.[62,63] A phase II trial incorporated direct vaccination with the PEPvIII fragment in 18 patients with newly diagnosed, EGFRvIII+ GBMs. Compared with historical matched controls, PFS was improved from 6.3 to 14.2 months.[64] A subsequent phase II trial treated 21 patients with the same histology using monthly intradermal vaccinations with concurrent temozolomide. The PFS was reported as 16.6 months, compared with the PFS in historical controls of 14.3 months.[65]

Despite encouraging preliminary results in these clinical trials, the true clinical efficacy of vaccine therapy is highly questionable. Peptide vaccines have been studied far more extensively in systemic cancers such as melanoma and prostate, breast, cervical, and colorectal cancers. In a rigorous review of these clinical trials, Rosenberg *et al.* reported only a 4% objective response rate, all of which occurred in melanoma trials exclusively.[66] They argue that the use of surrogate endpoints such as tumor necrosis or tumor lymphocyte infiltration do not adequately describe true effectiveness. Likewise, actual MRI responses are rarely seen in vaccine trials of recurrent glioma. Advocates of vaccines in gliomas have pointed to the longer PFS and OS seen in patients treated in the small phase I/II vaccine trials completed to date, compared with that seen in historical controls. The almost universal problem in interpreting these data, however, is that the patients treated on these vaccine trials tend to be highly selected, with very favorable prognostic factors, compared with the nonselected historical controls, making comparisons problematic.[67]

There are a number of reasons for skepticism regarding the ability of the current generation of crude vaccines to be clinically effective. Gliomas are poorly antigenic and they have an innate ability to induce immunosuppression mediated by regulatory T cells (Tregs). These factors are thought to be the major impediments to the success of immunotherapy. Glioma-driven immunosuppressive factors include transforming growth factor β, interleukin 10 (IL-10), and prostagladin E2 (PGE2).[68,69] Tregs expressing CD25 are increased in GBMs and may be pivotal in hindering immune responses. Thus, there is evidence that the local microenvironment of these tumors does not support a robust immune activation response. Bai *et al.* demonstrated that paradoxically, T cells may be simultaneously activated and tolerized in microenvironments of the same tumor.[70] Strategies to overcome glioma-induced immunosuppression and enhance vaccine efficacy include targeting immunosuppressive factors such as transforming growth factor β.[71]

STEM CELL THERAPY

Stem cells are characterized by the ability for extensive self-renewal and differentiation into diverse lineages. In contrast to the historically accepted stochastic model of cancer development and progression, the cancer stem cell model hypothesizes that most cancer cells within a tumor are not clonogenic and that only a small subpopulation of cells within the tumor truly possesses the capability to self-renew, proliferate, and (potentially) differentiate, as is seen in normal stem cells. Such tumor/cancer stem cells have been implicated as the principal governors of tumor initiation, growth, and resistance to treatment.[72–74]

Glioma stem cells were first isolated using the stem cell marker CD133[74,75]; however, more recent data has shown that the early neural stem cell marker SSEA-1 (CD15) is a more reliable marker of glioma stem cells.[76] Like normal neural stem cells (NSCs), cancer stem cells rely on well-orchestrated interactions within the microenvironment in order to flourish. Crucial signaling pathways for NSCs such as PI3K, notch, and Wnt are active in glioma stem cells as well,[77] and represent potential novel therapeutic targets.

In contrast to the historically accepted stochastic model of cancer development and progression, the cancer stem cell model hypothesizes that most cancer cells within a tumor are not clonogenic and that only a small subpopulation of cells within the tumor truly possesses the capability to self-renew, proliferate, and (potentially) differentiate, as is seen in normal stem cells. Such tumor/cancer stem cells have been implicated as the principal governors of tumor initiation, growth and resistance to treatment.

Notch receptor proteins are found on cell membranes, and once activated the intracellular portion is cleaved by γ-secretases, a class of integral membrane proteins, freeing it to associate with transcription factors that activate several genes essential for maintenance and renewal of neural stem cells Abnormal notch signaling has been implicated in glioma growth and progression,[78,79] and animal studies show blockade through inhibition of γ-secretases can restrict stem cell activity in vitro and tumor formation in vivo.[80,81] Notch inhibition with γ-secretase inhibitors is a new strategy under evaluation in early phase I clinical trials.

Sonic hedgehog homolog (SHH) is a crucial regulator of neurogenesis and the propagation of neural stem cells in the cerebellum.[82–84] SHH signaling has also been shown to regulate growth of medulloblastomas and gliomas,[85,86] and inhibition of this pathway may lead to reductions in glioblastoma stem-like cells, and increased responsiveness to therapy.[87] Although there is no controversy regarding the role that SHH plays in the biology of medulloblastomas, its role in gliomas remains controversial. The clinical development of SHH inhibitors, such as GDC-0449, is currently in phase I and II trials for medulloblastomas and other solid tumors where SHH appears to play an important role (ie, prostate cancer).

Although the identification of cancer stem cells has shifted our paradigm of oncogenesis and broadened the scope of therapeutic considerations, understanding their role in gliomagenesis is still in its relative infancy. Current research is ongoing in identifying optimal stem cell markers, tracking these cells in vivo, and sparing normal NSCs from the effects of targeted and cytotoxic therapies.

CONCLUSION

In recent years, basic and preclinical studies have revealed multiple new mechanisms of gliomagenesis and corresponding targets for treatment. Yet for the most part these discoveries have not translated to impressive improvements in response or survival for patients. Although this impasse is reached in many other disciplines of medicine, the challenges of conducting translational research for brain tumors are distinct. To begin with, the CNS circulation is essentially separated from the rest of the body by the BBB; drug delivery is therefore unpredictable and unreliable. Ideally we would analyze tumor tissue in order to confirm adequate concentrations of our agents; however, the difficulties of accessing brain tumor samples compared with other tumor samples in systemic cancers make this a difficult task. This also impedes our validation of biological effects in vivo, a crucial step in the assessment of new interventions and therapies. Finally, the validity of our endpoints for malignant glioma clinical trials is uncertain. The question of whether PFS or OS is the more appropriate endpoint for malignant glioma trials has long been debated. Moreover, the use of antiangiogenic therapies has changed patterns of enhancement and recurrence on imaging. Our current standard of radiographic response, Macdonald's criteria,[88] measures the enhancing tumor volume in 2 dimensions, and has proven inadequate to describe the changes seen with these newer therapies.

Nevertheless, much is being learned from the translational research now being undertaken to connect these basic science discoveries to clinical trials. Mechanisms of treatment resistance are better appreciated, and it has become increasingly apparent that inhibition of a single pathway or target is unlikely to affect tumor growth to the degree where durable clinical benefit is attained. Directed therapy against multiple pathways and growth factors is likely necessary, and with the prodigious number of potential targets available, the challenge will be in ascertaining the optimal combinations of drugs for individual patients. Parallel to this is the task of determining the unique molecular and genetic profiles of individual patient tumors so that the most effective therapies can be selected for those individuals. Developments in other arenas of brain tumor diagnosis and treatment such as neuroimaging, surgery, and radiation therapy are also evolving. Finally, great progress has been made in uncovering the molecular and genetic constitution of these tumors. Collectively, these advances have laid the foundation for significant improvements in treatments and outcomes for brain tumor patients.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. CBTRUS (2010). *CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2004–2006*. Hinsdale, IL: Central Brain Tumor Registry of the United States; February 2010. http://www.cbtrus.org.
2. Louis DN, Ohgaki H, Wiestler OD, et al., eds. *WHO Classification of Tumours of the Central Nervous System*. 3rd ed. Lyon, France: International Agency for Research on Cancer; 2007.
3. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* 2005; 352: 987–996.
4. Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol* 2009; 10: 459–966.
5. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 2008; 455: 1061–1068.
6. Verhaak RG, Hoadley KA, Purdom E, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* 2010; 17: 98–110.
7. Li A, Walling J, Ahn S, et al. Unsupervised analysis of transcriptomic profiles reveals six glioma subtypes. *Cancer Res* 2009; 69: 2091–2099.
8. Madhavan S, Zenklusen JC, Kotliarov Y, et al. Rembrandt: helping personalized medicine become a reality through integrative translational research. *Mol Cancer Res* 2009; 7: 157–167.
9. Belanich M, Pastor M, Randall T, et al. Retrospective study of the correlation between the DNA repair protein alkyltransferase and survival of brain tumor patients treated with carmustine. *Cancer Res* 1996; 56: 783–788.

10. Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. *Nat Rev Cancer* 2004; 4: 296–307.
11. Jaeckle KA, Eyre HJ, Townsend JJ, et al. Correlation of tumor O6 methylguanine-DNA methyltransferase levels with survival of malignant astrocytoma patients treated with bis-chloroethylnitrosourea: a Southwest Oncology Group study. *J Clin Oncol* 1998; 16: 3310–3315.
12. Esteller M, Garcia-Foncillas J, Andion E, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents [published correction appears in *N Engl J Med* 2000;343:1740]. *N Engl J Med* 2000; 343: 1350–1354.
13. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 2005; 352: 997–1003.
14. Dolan ME, Mitchell RB, Mummert C, et al. Effect of O6-benzylguanine analogues on sensitivity of human tumor cells to the cytotoxic effects of alkylating agents. *Cancer Res* 1991; 51: 3367–3372.
15. Wedge SR, Porteous JK, Newlands ES. 3-aminobenzamide and/or O6-benzylguanine evaluated as an adjuvant to temozolomide or BCNU treatment in cell lines of variable mismatch repair status and O6-alkylguanine-DNA alkyltransferase activity. *Br J Cancer* 1996; 74: 1030–1036.
16. Quinn JA, Jiang SX, Reardon DA, et al. Phase II trial of temozolomide plus O6-benzylguanine in adults with recurrent, temozolomide-resistant malignant glioma. *J Clin Oncol* 2009; 27: 1262–1267.
17. Quinn JA, Jiang SX, Carter J, et al. Phase II trial of Gliadel plus O6-benzylguanine in adults with recurrent glioblastoma multiforme. *Clin Cancer Res* 2009; 15: 1064–1068.
18. Quinn JA, Pluda J, Dolan ME, et al. Phase II trial of carmustine plus O(6)-benzylguanine for patients with nitrosourea-resistant recurrent or progressive malignant glioma. *J Clin Oncol* 2002; 20: 2277–2283.
19. Tolcher AW, Gerson SL, Denis L, et al. Marked inactivation of O6-alkylguanine-DNA alkyltransferase activity with protracted temozolomide schedules. *Br J Cancer* 2003; 88: 1004–1011.
20. Clarke JL, Iwamoto FM, Sul J, et al. Randomized phase II trial of chemoradiotherapy followed by either dose-dense or metronomic temozolomide for newly diagnosed glioblastoma. *J Clin Oncol* 2009; 27: 3861–3867.
21. Libermann TA, Nusbaum HR, Razon N, et al. Amplification, enhanced expression and possible rearrangement of EGF receptor gene in primary human brain tumours of glial origin. *Nature* 1985; 313: 144–147.
22. Yamazaki H, Fukui Y, Ueyama Y, et al. Amplification of the structurally and functionally altered epidermal growth factor receptor gene (c-erbB) in human brain tumors. *Mol Cell Biol* 1988; 8: 1816–1820.
23. Brown PD, Krishnan S, Sarkaria JN, et al. Phase I/II trial of erlotinib and temozolomide with radiation therapy in the treatment of newly diagnosed glioblastoma multiforme: North Central Cancer Treatment Group Study N0177. *J Clin Oncol* 2008; 26: 5603–5609.
24. Furnari FB, Fenton T, Bachoo RM, et al. Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes Dev* 2007; 21: 2683–2710.
25. Sugawa N, Ekstrand AJ, James CD, et al. Identical splicing of aberrant epidermal growth factor receptor transcripts from amplified rearranged genes in human glioblastomas. *Proc Natl Acad Sci U S A* 1990; 87: 8602–8606.
26. Rich JN, Reardon DA, Peery T, et al. Phase II trial of gefitinib in recurrent glioblastoma. *J Clin Oncol* 2004; 22: 133–142.
27. De Witt Hamer PC. Small molecule kinase inhibitors in glioblastoma: a systematic review of clinical studies. *Neuro Oncol* 2010; 12: 304–316.
28. Mellinghoff IK, Wang MY, Vivanco I, et al. Molecular determinants of the response of glioblastomas to EGFR kinase inhibitors. *N Engl J Med* 2005; 353: 2012–2024.
29. Cloughesy TF, Yoshimoto K, Nghiemphu P, et al. Antitumor activity of rapamycin in a Phase I trial for patients with recurrent PTEN-deficient glioblastoma. *PLoS Med* 2008; 5: e8.
30. Stommel JM, Kimmelman AC, Ying H, et al. Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies. *Science* 2007; 318: 287–290.
31. Folkman J, Merler E, Abernathy C, et al. Isolation of a tumor factor responsible for angiogenesis. *J Exp Med* 1971; 133: 275–288.
32. Gimbrone MA Jr, Leapman SB, Cotran RS, et al. Tumor dormancy in vivo by prevention of neovascularization. *J Exp Med* 1972; 136: 261–276.
33. Folkman J. Tumor angiogenesis: therapeutic implications. *N Engl J Med* 1971; 285: 1182–1186.
34. Keck PJ, Hauser SD, Krivi G, et al. Vascular permeability factor, an endothelial cell mitogen related to PDGF. *Science* 1989; 246: 1309–1312.
35. Jain RK, di Tomaso E, Duda DG, et al. Angiogenesis in brain tumours. *Nat Rev Neurosci* 2007; 8: 610–622.
36. Sun L, Hui AM, Su Q, et al. Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain. *Cancer Cell* 2006; 9: 287–300.
37. Jain RK. Normalizing tumor vasculature with anti-angiogenic therapy: a new paradigm for combination therapy. *Nat Med* 2001; 7: 987–989.
38. Batchelor TT, Sorensen AG, di Tomaso E, et al. AZD2171, a pan-VEGF receptor tyrosine kinase inhibitor, normalizes tumor vasculature and alleviates edema in glioblastoma patients. *Cancer Cell* 2007; 11: 83–95.
39. Batchelor TT, Duda DG, di Tomaso E, et al. Phase II study of cediranib, an oral pan-vascular endothelial growth factor receptor tyrosine kinase inhibitor, in patients with recurrent glioblastoma. *J Clin Oncol* 2010; 28: 2817–2823.
40. Stark-Vance V. Bevacizumab and CPT-11 in the treatment of relapsed malignant glioma. *Neuro Oncol* 2005; 7: 369.
41. Friedman HS, Prados MD, Wen PY, et al. Bevacizumab alone and in combination with irinotecan in recurrent glioblastoma. *J Clin Oncol* 2009; 27: 4733–4740.
42. Kreisl TN, Kim L, Moore K, et al. Phase II trial of single-agent bevacizumab followed by bevacizumab plus irinotecan at tumor progression in recurrent glioblastoma. *J Clin Oncol* 2009; 27: 740–745.
43. Vredenburgh JJ, Desjardins A, Herndon JE II, et al. Bevacizumab plus irinotecan in recurrent glioblastoma multiforme. *J Clin Oncol* 2007; 25: 4722–4729.
44. Guo P, Hu B, Gu W, et al. Platelet-derived growth factor-B enhances glioma angiogenesis by stimulating vascular endothelial growth factor expression in tumor endothelia and by promoting pericyte recruitment. *Am J Pathol* 2003; 162: 1083–1093.

45. Leenders WP, Küsters B, Verrijp K, et al. Antiangiogenic therapy of cerebral melanoma metastases results in sustained tumor progression via vessel co-option. *Clin Cancer Res* 2004; 10: 6222–6230.
46. Lamszus K, Kunkel P, Westphal M. Invasion as limitation to anti-angiogenic glioma therapy. *Acta Neurochir Suppl* 2003; 88: 169–177.
47. Kim KJ, Li B, Winer J, et al. Inhibition of vascular endothelial growth factor–induced angiogenesis suppresses tumour growth in vivo. *Nature* 1993; 362: 841–844.
48. Rubenstein JL, Kim J, Ozawa T, et al. Anti-VEGF antibody treatment of glioblastoma prolongs survival but results in increased vascular cooption. *Neoplasia* 2000; 2: 306–314.
49. Carson MJ, Reilly CR, Sutcliffe JG, et al. Mature microglia resemble immature antigen-presenting cells. *Glia* 1998; 22: 72–85.
50. Aloisi F. Immune function of microglia. *Glia* 2001; 36: 165–179.
51. Fathallah-Shaykh HM, Gao W, Cho M, et al. Priming in the brain, an immunologically privileged organ, elicits anti-tumor immunity. *Int J Cancer* 1998; 75: 266–276.
52. Ni HT, Spellman SR, Jean WC, et al. Immunization with dendritic cells pulsed with tumor extract increases survival of mice bearing intracranial gliomas. *J Neurooncol* 2001; 51: 1–9.
53. Yu JS, Wheeler CJ, Zeltzer PM, et al. Vaccination of malignant glioma patients with peptide-pulsed dendritic cells elicits systemic cytotoxicity and intracranial T-cell infiltration. *Cancer Res* 2001; 61: 842–847.
54. Kikuchi T, Akasaki Y, Irie M, et al. Results of a phase I clinical trial of vaccination of glioma patients with fusions of dendritic and glioma cells. *Cancer Immunol Immunother* 2001; 50: 337–344.
55. Liau LM, Prins RM, Kiertscher SM, et al. Dendritic cell vaccination in glioblastoma patients induces systemic and intracranial T-cell responses modulated by the local central nervous system tumor microenvironment. *Clin Cancer Res* 2005; 11: 5515–5525.
56. Suter T, Biollaz G, Gatto D, et al. The brain as an immune privileged site: dendritic cells of the central nervous system inhibit T cell activation. *Eur J Immunol* 2003; 33: 2998–3006.
57. Rosenberg SA, Sherry RM, Morton KE, et al. Tumor progression can occur despite the induction of very high levels of self tumor antigen-specific CD8+ T cells in patients with melanoma. *J Immunol* 2005; 175: 6169–6176.
58. Kikuchi T, Akasaki Y, Abe T, et al. Vaccination of glioma patients with fusions of dendritic and glioma cells and recombinant human interleukin 12. *J Immunother* 2004; 27: 452–459.
59. Walker DG, Laherty R, Tomlinson FH, et al. Results of a phase I dendritic cell vaccine trial for malignant astrocytoma: potential interaction with adjuvant chemotherapy. *J Clin Neurosci* 2008; 15: 114–121.
60. Ludewig B, Ochsenbein AF, Odermatt B, et al. Immunotherapy with dendritic cells directed against tumor antigens shared with normal host cells results in severe autoimmune disease. *J Exp Med* 2000; 191: 795–804.
61. Sampson JH, Archer GE, Mitchell DA, et al. Tumor-specific immunotherapy targeting the EGFRvIII mutation in patients with malignant glioma. *Semin Immunol* 2008; 20: 267–275.
62. Heimberger AB, Crotty LE, Archer GE, et al. Epidermal growth factor receptor VIII peptide vaccination is efficacious against established intracerebral tumors. *Clin Cancer Res* 2003; 9: 4247–4254.
63. Wu AH, Xiao J, Anker L, et al. Identification of EGFRvIII-derived CTL epitopes restricted by HLA A0201 for dendritic cell based immunotherapy of gliomas. *J Neurooncol* 2006; 76: 23–30.
64. Heimberger AB, Hussain SF, Aldape K. Tumor specific peptide vaccination in newly-diagnosed patients with GBM. *J Clin Oncol* 2006; 24: 2529.
65. Sampson JH, Archer GE, Bigner DD. Effect of EGFRvIII-targeted vaccine (CDX-110) on immune response and TTP when given with simultaneous standard and continuous temozolomide in patietns with GBM. *J Clin Oncol* 2008; 26: 2011.
66. Rosenberg SA, Yang JC, Restifo NP. Cancer immunotherapy: moving beyond current vaccines. *Nat Med* 2004; 10: 909–915.
67. Fine HA. Toward a glioblastoma vaccine: promise and potential pitfalls. *J Clin Oncol* 2004; 22: 4240–4243.
68. Wrzesinski SH, Wan YY, Flavell RA. Transforming growth factor-beta and the immune response: implications for anticancer therapy. *Clin Cancer Res* 2007; 13: 5262–5270.
69. O'Garra A, Barrat FJ, Castro AG, et al. Strategies for use of IL-10 or its antagonists in human disease. *Immunol Rev* 2008; 223: 114–131.
70. Bai A, Higham E, Eisen HN, et al. Rapid tolerization of virus-activated tumor-specific CD8+ T cells in prostate tumors of TRAMP mice. *Proc Natl Acad Sci U S A* 2008; 105: 13003–13008.
71. Uhl M, Aulwurm S, Wischhusen J, et al. SD-208, a novel transforming growth factor beta receptor I kinase inhibitor, inhibits growth and invasiveness and enhances immunogenicity of murine and human glioma cells in vitro and in vivo. *Cancer Res* 2004; 64: 7954–7961.
72. Bao S, Wu Q, McLendon RE, et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* 2006; 444: 756–760.
73. Dietrich J, Imitola J, Kesari S. Mechanisms of disease: the role of stem cells in the biology and treatment of gliomas. *Nat Clin Pract Oncol* 2008; 5: 393–404.
74. Singh SK, Hawkins C, Clarke ID, et al. Identification of human brain tumour initiating cells. *Nature* 2004; 432: 396–401.
75. Singh SK, Clarke ID, Terasaki M, et al. Identification of a cancer stem cell in human brain tumors. *Cancer Res* 2003; 63: 5821–5828.
76. Son MJ, Woolard K, Nam DH, et al. SSEA-1 is an enrichment marker for tumor-initiating cells in human glioblastoma. *Cell Stem Cell* 2009; 4: 440–452.
77. Phillips HS, Kharbanda S, Chen R, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* 2006; 9: 157–173.
78. Purow BW, Haque RM, Noel MW, et al. Expression of Notch-1 and its ligands, Delta-like-1 and Jagged-1, is critical for glioma cell survival and proliferation. *Cancer Res* 2005; 65: 2353–2363.
79. Kanamori M, Kawaguchi T, Nigro JM, et al. Contribution of Notch signaling activation to human

glioblastoma multiforme. *J Neurosurg* 2007; 106: 417–427.

80. Fan X, Matsui W, Khaki L, et al. Notch pathway inhibition depletes stem-like cells and blocks engraftment in embryonal brain tumors. *Cancer Res* 2006; 66: 7445–7452.
81. Fan X, Mikolaenko I, Elhassan I, et al. Notch1 and notch2 have opposite effects on embryonal brain tumor growth. *Cancer Res* 2004; 64: 7787–7793.
82. Machold R, Hayashi S, Rutlin M, et al. Sonic hedgehog is required for progenitor cell maintenance in telencephalic stem cell niches. *Neuron* 2003; 39: 937–950.
83. Lai K, Kaspar BK, Gage FH, et al. Sonic hedgehog regulates adult neural progenitor proliferation in vitro and in vivo. *Nat Neurosci* 2003; 6: 21–27.
84. Cai C, Thorne J, Grabel L. Hedgehog serves as a mitogen and survival factor during embryonic stem cell neurogenesis. *Stem Cells* 2008; 26: 1097–1108.
85. Goodrich LV, Milenkovic L, Higgins KM, et al. Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science* 1997; 277: 1109–1113.
86. Clement V, Sanchez P, de Tribolet N, et al. HEDGEHOG-GLI1 signaling regulates human glioma growth, cancer stem cell self-renewal, and tumorigenicity. *Curr Biol* 2007; 17: 165–172.
87. Bar EE, Chaudhry A, Lin A, et al. Cyclopamine-mediated hedgehog pathway inhibition depletes stem-like cancer cells in glioblastoma. *Stem Cells* 2007; 25: 2524–2533.
88. Macdonald DR, Cascino TL, Schold SC Jr, et al. Response criteria for phase II studies of supratentorial malignant glioma. *J Clin Oncol* 1990; 8: 1277–1280.
89. Lieberman FS, Cloughesy T, Fine HA, et al. NABTC phase I/II trial of ZD-1839 for recurrent malignant gliomas and unresectable meningiomas. *J Clin Oncol* 2004; 22(14S): Abstract 1510.
90. Uhm JH, Ballman KV, Giannini C, et al. Phase II study of ZD1839 in patients with newly diagnosed grade 4 astrocytoma. *J Clin Oncol* 2004; 22(14S): Abstract 1505.
91. Franceschi E, Cavallo G, Lonardi S, et al. Gefitinib in patients with progressive high-grade gliomas: a multicentre phase II study by Gruppo Italiano Cooperativo di Neuro-Oncologia (GICNO). *Br J Cancer* 2007; 96: 1047–1051.
92. Vogelbaum MA, Peereboom DM, Stevens G, et al. Phase II study of erlotinib single agent therapy in recurrent glioblastoma multiforme. Paper presented at: 29th ESMO Congress; October 29–November 3, 2004; Vienna, Austria.
93. Raizer JJ, Abrey LE, Wen P, et al. A phase II trial of erlotinib (OSI-774) in patients (pts) with recurrent malignant gliomas (MG) not on EIAEDs. *J Clin Oncol* 2004; 22(14S): Abstract 1502.
94. Cloughesy T, Yung A, Vredenburgh J, et al. Phase II study of erlotinib in recurrent GBM: molecular predictors of outcome. *J Clin Oncol* 2005; 23(16S, part 1): Abstract 1507.
95. Van den Bent MJ, Brandes AA, Rampling R, et al. Randomized phase II trial of erlotinib versus temozolomide or carmustine in recurrent glioblastoma: EORTC brain tumor group study 26034. *J Clin Oncol* 2009; 27: 1268–1274.

Myeloproliferative Neoplasms: New Translational Therapies

John Mascarenhas, MD, and Ronald Hoffman, MD

Mount Sinai School of Medicine, New York, NY

OUTLINE

History of Myeloproliferative Neoplasms
Classification of Myeloproliferative Neoplasms
  Chronic Myelogenous Leukemia
  Polycythemia Vera and Essential Thrombocythemia
  Primary Myelofibrosis
Translational Medicine in Myeloproliferative Neoplasms
  Chronic Myelogenous Leukemia
  Philadelphia Chromosome Negative Myeloproliferative Neoplasms
Experimental Trials at Mount Sinai School of Medicine
Conclusion

ABSTRACT

The myeloproliferative neoplasms represent a diverse group of hematologic malignancies that have been the subject of intense investigation over the last decade. Although clinical trials of the much anticipated small molecule inhibitors of Janus kinase 2 have shown that these experimental agents are successful in palliating many of the symptoms associated with the myeloproliferative neoplasms, they have not been reported to affect the disease initiating hematopoietic stem cell population or to alter the natural history of these disorders. Investigators remain optimistic that new information about the genetic and cellular origins gained from the efforts of numerous laboratories will ultimately translate in to the identification of new drug targets and more effective therapies. We hypothesize that ultimately, the use of combinations of drugs including chromatin modifying agents, immunomodulatory agents, anti-apoptotic agents, cellular therapies and monoclonal antibodies will be required to effectively treat patients with myeloproliferative neoplasms. ***Mt Sinai J Med 77:667–683, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: BCR-ABL, chronic myelogenous leukemia, clinical trials, essential thrombocythemia, JAK2 mutation, myeloproliferative neoplasms, polycythemia vera, primary myelofibrosis, stem cell transplantation, tyrosine kinase inhibitors.

Perhaps in no other area of malignant hematology has the development of novel experimental therapies based on rational scientific understanding of molecular pathogenesis been as rewarding as in the area of myeloproliferative neoplasms (MPNs). In the last 5 years, a new explosion of advances in therapeutic options based on molecular insights has invigorated the field. In 2008, the World Health Organization revised the classification of the hematopoietic malignancies listing within the newly termed category MPN (formerly called myeloproliferative disorders). These disorders include chronic myelogenous leukemia (CML), polycythemia vera (PV), essential thrombocythemia (ET), primary myelofibrosis (PMF), systemic mastocytosis, chronic eosinophilic leukemia not otherwise specified, chronic neutrophilic leukemia, and

Perhaps in no other area of malignant hematology has the rational scientific understanding of molecular pathogenesis facilitated the development of novel experimental therapies as in the area of myeloproliferative neoplasms (MPNs) during the last 5 years.

Address Correspondence to:
John Mascarenhas
Division of
Hematology/Oncology
Box 1079
Mount Sinai School of Medicine
New York, NY
Email: john.mascarenhas@mssm.edu

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20225

© 2010 Mount Sinai School of Medicine

MPN unclassifiable.[1] Chronic myelogenous leukemia remains unique among this group due to its established monogenetic chromosomal translocation (Philadelphia [Ph] chromosome). This cytogenetic translocation is invariably present and its contribution to disease pathogenesis has been well documented. This review will discuss both the historic and current scientific advances in MPNs. Chronic myelogenous leukemia and the classic Ph-negative MPNs–PV, ET, and PMF–will be the major focus, and we will exclude the other MPNs due to their relative low incidence.

HISTORY OF MYELOPROLIFERATIVE NEOPLASMS

Chronic myelogenous leukemia was first described in 1845 by Dr John Hughes Bennett in the *Edinburgh Medical and Surgical Journal* as a case of spleen hypertrophy leading to death from "suppuration of the blood." [2] In the 1960s, Nowell and Hungerford were the first to recognize the presence of a persistent and reproducible chromosomal abnormality in the cells from the blood and bone marrow of patients with CML and postulated the role of genetic mutations in the origin of leukemia.[3] As a result of the work by Janet Rowley in the 1970s, it was revealed that this Philadelphia chromosome (so named because of the city in which it was discovered by Nowell and Hungerford) was the result of a reciprocal translocation between chromosomes 9 and 22: t(9;22)(q34;q11).[4] In the 1980s, the t(9:22) chromosomal abnormality was shown to result in the production of a fusion oncogene bringing into juxtaposition the breakpoint cluster region gene from chromosome 22 (region q11) with the Abl1 gene on chromosome 9 (region q34).[5–7]

Polycythemia vera was the first MPN described by Hippocrates in a patient with noted plethora and first reported in the medical literature by Luis Henry Vaquez in 1892 as a case describing a man with cyanosis and an elevated hemoglobin.[8] In 1903, William Osler reported additional cases with a similar clinical presentation.[9] It was not until 1951 that William Dameshek (cofounder of the journal *Blood* and the American Society of Hematology) grouped these disorders under a unifying theme of myeloproliferation and called them collectively the myeloproliferative disorders.[10] In 1967, Dr Louis Wasserman founded the Polycythemia Vera Study Group (PVSG) with funding from the National Cancer Institute with a mission to better characterize the natural history of PV and conduct clinical trials and formulate standard diagnostic criteria.[11] Perhaps the most significant contributions to the treatment paradigm of PV that resulted from the PVSG studies were from the randomized PVSG-01 trial, in which phlebotomy alone was shown to be superior to treatment with chlorambucil or radioactive phosphorus (P-32) and the nonrandomized PVSG-08 showing the superiority of hydroxyurea as a cytoreductive agent in thrombotic risk reduction.[12,13]

Mount Sinai School of Medicine (MSSM) has played a very important role in the history of MPN research and treatment. From the inception of the PVSG by Dr Wasserman (for which the current malignant hematology house staff service is named) to the current Myeloproliferative Disorder program headed by Dr Ronald Hoffman, MSSM has been at the forefront of advances in the understanding of the pathobiology and treatment of MPNs. Much pride remains at MSSM for the contributions that have been and continue to be made in the MPNs and include the past and current efforts of researchers including Steven Fruchtman, Vesna Najfeld, Harriet Gilbert, Paul Berk, Neal Weinreb, Esmail Zanjani, John Lutton, Nathaniel Wisch, Rona Weinberg, Mingjiang Xu, Anna Rita Migliaccio, and Ronald Hoffman.

CLASSIFICATION OF MYELOPROLIFERATIVE NEOPLASMS

The MPNs, as a class, are characterized by a hyperproliferative bone marrow, splenomegaly, constitutional symptoms, and a propensity for thrombosis. Thrombocytosis alone is seen in ET, and an absolute erythrocytosis often accompanied by leukocytosis and thrombocytosis is characteristic of PV. CML is represented by an absolute granulocytic hyperplasia with full cellular maturation. Patients with PMF can have either leukocytosis or leukopenia, thrombocytosis, or thrombocytopenia and most often develop anemia. Both PV and ET can transform to MF and are denoted as post-PV MF and post-ET MF, respectively.[14] Due to significant overlapping phenotypic and pathogenetic characteristics, a combination of clinical, bone marrow morphological, immunohistochemical, cytogenetic, and genetic features are required to better define each MPN.[15] Commonalities within the MPNs

The MPNs, as a class, are characterized by a hyperproliferative bone marrow, splenomegaly, constitutional symptoms, and a propensity for thrombosis.

DOI:10.1002/MSJ

include a hematopoietic stem cell of origin, a hyperproliferative marrow composed of mature myeloid cells, and the possibility for one MPN to transform into another and ultimately into a treatment-refractory secondary acute leukemia, usually of a myeloid phenotype. It is based on these discrete differences that clinicians ascribe a disease name and prescribe a treatment plan for patients suffering from these malignancies. Although the MPNs are chronic by definition, they are often progressive by nature and are associated with a risk for leukemic transformation that is not always obvious at time of diagnosis.

Chronic Myelogenous Leukemia

Chronic myelogenous leukemia represents 13%–20% of all new leukemia diagnoses in the United States annually.[16,17] Usually, but not always, CML progresses through 3 defined phases, termed the chronic phase (CP), accelerated phase (AP), and blast phase (BP). Approximately 90% of patients are diagnosed and initiated on treatment in the CP. The median age at diagnosis is 53 and the most common symptoms include fatigue, anorexia, and weight loss, although 40% of patients are asymptomatic at presentation.[17] Splenomegaly is found in 50% of patients and an elevated granulocyte count and platelet count are characteristic on the peripheral smear.[18] Untreated, CML-CP can persist for 3 to 6 years and will ultimately progress to fatal CML-BP, with the majority of patients transitioning through CML-AP.[19] Prior to the advent of tyrosine kinase inhibitor (TKI) therapy, CML was treated with chemotherapeutic agents such as busulfan and hydroxyurea that only promised to modify hematologic parameters and offered no true alteration in disease course.[20] Meta-analyses have clearly shown a statistically significant improvement in survival with interferon-alfa (IFN-alfa) therapy over such chemotherapy regimens,[21] and IFN-alfa was the accepted first-line therapy for CML before imatinib. The revolutionary success of targeted therapy in the form of TKIs in CML will be discussed below.

Polycythemia Vera and Essential Thrombocythemia

Patients with ET can expect a similar median survival as age-matched cohorts.[22–24] Patients with PV, however, have been found to have a 1.6-fold higher mortality than control populations.[25] Arterial and venous thrombosis, hemorrhage, evolution to MF, and leukemic transformation are the main causes of morbidity and mortality. Patients with PV/ET are stratified by risk for thrombosis by age (>60 years) and history of thrombotic events.[26] Those PV/ET patients at high risk are treated with cytoreductive agents such as hydroxyurea with the goal to decrease subsequent thrombotic events (myocardial infarction and cerebrovascular accidents are the most likely to result in death in the older age group).

Primary Myelofibrosis

Patients with PMF have a median overall survival of 5 years. Several risk-stratification systems have been created as prognostic tools and are used within a research context when considering an individual for treatment on an experimental protocol.[24,27,28] Survival in MF is limited mainly by leukemic transformation and complications arising from progressive splenomegaly and portal hypertension.[27] Due to the progressive nature of MF and the poor outcome of the majority of patients, great emphasis and attention has been placed on novel agents being tested in ongoing phase I, II, and III clinical studies. In the past, many therapies have been investigated including traditional chemotherapy, splenic radiation, splenectomy, and stem cell transplantation (SCT).[29] Therapeutic approaches to the treatment of MF have included immunomodulatory agents such as thalidomide and lenalidomide in combination with prednisone; these combinations offer response rates ranging from 30% to 50% and are associated with toxicities including neuropathy, sedation, constipation, myelosuppression, and a potential increased thrombotic risk.[30–33] Androgens have also been used to selectively manage the anemia associated with MF, with response rates close to 40%.[34] Erythropoiesis-stimulating agents have been examined in prospective studies with conflicting results and are not approved by the US Food and Drug Administration for use in MPNs.[35–38] Chemotherapeutics including hydroxyurea, melphalan, busulfan, and 2-CdA are used to control the myeloproliferative aspect of MF,[39–42] but they have failed to show any significant impact on disease progression and likely increase the risk for leukemic transformation.

Hematopoietic SCT is the only treatment modality for MF that offers the potential for cure at a recognized risk of transplant-related mortality (TRM) of approximately 10% and 30% for reduced-intensity conditioning (RIC) and myeloablative regimens, respectively.[43,44] Stem cell transplantation is an appropriate consideration for those patients aged <65 years with a good performance status and the absence of significant competing comorbidities with a 10/10 HLA-matched donor. Current clinical trials are defining the role of RIC SCT in MF within the Myeloproliferative Disorders Research Consortium (MPD-RC), and interim results appear favorable for matched related donors with a 79% overall response rate by International Working Group for Myelofibrosis Research and Treatment criteria and a TRM of 13%.[45]

TRANSLATIONAL MEDICINE IN MYELOPROLIFERATIVE NEOPLASMS

Chronic Myelogenous Leukemia

True success in translation of scientific discovery into a novel effective therapeutic was first seen in CML with the development of imatinib mesylate (Gleevec). The story of imatinib in CML is the best example of targeted drug development based on new understanding of a molecular pathogenetic mechanism. This process, however, did not happen overnight; many decades of scientific discovery and drug development led to the discovery and clinical development of imatinib. In the late 1980s, Brian Druker collaborated with others in academia and the pharmaceutical industry on a drug-discovery program that would result in the start of clinical trials in 1998 with imatinib versus the then standard of care, IFN + low-dose cytarabine. The International Randomized Study of Interferon and STI571 (IRIS) ultimately culminated in FDA approval for the use of imatinib as front-line therapy for patients with newly diagnosed CML.[46]

> *The story of imatinib in chronic myelogenous leukemia is the best example of targeted drug development based on new understanding of a molecular pathogenetic mechanism.*

The BCR-ABL-1 oncogene product gives rise to a constitutively activated and unregulated cytoplasmic tyrosine kinase that causes uncontrolled proliferation and impaired differentiation of CML cells.[47] Proposed mechanisms of BCR-ABL–mediated malignant transformation are altered adhesion cell properties, activation of mitogenic signaling, and inhibition of apoptosis.[47] ABL tyrosine kinase dysregulation results in the activation of a multitude of signaling pathways including signal transducer and activator of transcription (STAT), MYC, mitogen-activated protein kinase (MAPK), RAS, and phosphoinositide 3-kinase (PI3K), which induce gene activation promoting cell survival, differentiation, and avoidance of apoptosis. In the IRIS trial that led to approval of imatinib for the treatment of CML, imatinib was shown to be superior in conferring a complete cytogenetic remission in 76% of patients versus 15% on the IFN and low-dose cytarabine arm, and after 8 years of follow-up continues to boast an overall survival of 86%.[46,48] Dasatinib (Sprycel) and nilotinib (Tasigna) are potent second-generation oral BCR-ABL1 inhibitors that have recently been shown to induce higher rates of early cytogenetic and even molecular responses in CML-CP and are increasingly being used as first treatment options at diagnosis.[49,50] In very recent published reports, both nilotinib and dasatinib have been compared in randomized phase III studies with imatinib in newly diagnosed patients with CML and have shown statistically significant improvements in achieving molecular and cytogenetic responses over imatinib.[51,52] Gastrointestinal complaints and fluid retention were the most common complaints with imatinib-treated patients; rash and headache were most associated with treatment with nilotinib; and pleural effusions were most notably attributed to treatment with dasatinib. It should be noted that discontinuation rates were similar between the 3 agents, and the rates of non-hematologic grade III/IV adverse events were low for each drug. It has previously been shown that a delay in achieving a complete cytogenetic response with imatinib predicts an eventual increased risk of disease progression.[53] An advantage of 3 available effective TKIs for the treatment of CML is in that each differs in their toxicity profile, and one may be better tolerated than another in a particular patient. Toxicities that can influence treatment decisions for a given patient include the hematological toxicity that is associated with each of the TKIs, gastrointestinal symptoms with imatinib, development of pleural effusions with dasatinib, and QTc interval prolongation with nilotinib.

> *In the International Randomized Study of Interferon and STI571 trial, imatinib was shown to be superior in conferring a complete cytogenetic remission of CML in 76% of patients versus 15% on the interferon-alfa and low-dose cytarabine arm. After 8 years of follow-up, it continues to boast an overall survival of 86%.*

Multiple mechanisms can lead to resistance of CML cells to TKI therapy. The development of

BCR-ABL1 kinase domain mutations is the major cause of resistance to imatinib therapy.[54,55] Point mutations in the kinase domain of the BCR-ABL1 gene result in conformational changes in the protein product that reduce the binding capacity of imatinib, and >100 such mutations have been identified.[56,57] Certain mutations such as T315I are resistant to all known TKIs and predict for poor survival. Not all mechanisms of resistance are BCR-ABL1–based, as is the case with low expression of the human organic cation transporter 1, which leads to reduced influx of imatinib into the CML cells.[58] The human organic cation transporter 1 assay is commercially available and can predict response to imatinib and identify patients who will require higher-dose therapy to overcome this mechanism of resistance.[59]

Despite the tremendous success of the current TKIs in inducing meaningful cytogenetic and molecular responses that translate to improved survival for patients with CML, TKI-resistant mutant BCR-ABL1 isoforms have emerged, and other experimental agents are being developed and tested to address this issue. Newer oral agents include bosutinib (SKI-606), a dual Src/Abl tyrosine kinase inhibitor which has shown activity against imatinib-resistant BCR-ABL1 mutant isoforms other than T3151 in chronic-phase CML.[60] Forty-six percent of imatinib-resistant patients achieved a complete cytogenetic remission on bosutinib after 14 months' follow-up. The intravenous multi-TKI XL228 targets, among other proteins, BCR-ABL1.[61] In a small phase I study, XL228 has been shown to have activity against the T315I isoform and offers a potential second-line agent for resistant CML patients with the T315I mutation who would otherwise require SCT. The malignant hematopoietic stem cell (HSC) in CML is able to survive TKI therapy despite complete cytogenetic/molecular responses, as is evidenced by relapse of disease after withdrawal of imatinib even after 2 years of a patient being in complete molecular remission.[62] Other classes of agents, including homoharringtonine (omacetaxine mepesuccinate), Aurora kinase inhibitors (PHA-739 358), and farnesyl transferase inhibitors (BMS-214 662), are being investigated in preclinical models as well as in early-phase clinical studies for their activity against the CML HSC and TKI-resistant BCR-ABL1 isoforms.[63–66] Lastly, immunotherapy with a novel BCR-ABL–derived peptide vaccine has been shown to elicit a BCR-ABL peptide-specific CD4+ T-cell response in a small phase II setting.[67] Although causality has not yet been definitely proved, cytogenetic and molecular responses have been observed in patients receiving vaccine therapy with concurrent IFN, imatinib, or in the setting of relapsed disease post-SCT. A recent report with long-term follow-up (>5 years) appears promising, with some patients with persistent disease detected with fluorescence in situ hybridization obtaining a complete cytogenetic remission, and even some patients obtaining complete molecular response with combination therapy.[68] Other small studies have been conducted exploiting an immunotherapeutic approach with vaccination and represent a novel mechanism to improve upon suboptimal responses in patient on TKIs.[69–71] These studies will need to be followed up on larger scale and are already being tested in multicenter trials in Europe.

Stem cell transplantation is still the only therapy that has clearly demonstrated the ability to cure patients with CML.[72–76] However, due to the enormous success of TKI therapy in CML, SCT is no longer considered a front-line therapeutic option and is reserved primarily for those CML patients who develop TKI-resistant disease.[77,78] There is however, no published prospective study demonstrating efficacy in this group, although case reports support the use of SCT for patients with the T3151 mutation.[79] In the 1990s, CML was the most common indication for SCT in the world, but after 1999 it declined substantially and is now an uncommon SCT indication at most US centers.[80] The main factors influencing transplant outcome in CML include CML stage at time of transplant (best in CP), donor type (sibling over unrelated), time interval from diagnosis (<12 months), age of recipient, and donor-recipient sex combination.[81] Preparative regimens have also evolved over the years, and improvements in TRM are reflected in retrospective studies. The busulfan/cyclophosphamide conditioning regimen has been shown in retrospective studies to achieve a 3-year survival of 86% with a risk of non-leukemic death of 14%.[82] Of 3018 CML patients from Europe undergoing SCT between 2000 and 2003, the 2-year survival was 61% and TRM was 30%.[81] The presence of chronic graft-versus-host disease (GVHD) has also been shown to be predictive of reduced risk of relapsed disease.[83] An important lesson regarding the correlation of GVHD and the phenomenon of a graft-versus-leukemia effect was learned in the setting of SCT and CML. Graft-versus-host disease is primarily mediated by complex interactions centering on the donor T lymphocyte populations reacting against recipient determinants. T lymphocyte–deplete SCTs offer a lower incidence of GVHD, leading to a lower TRM, but risk a higher relapse rate that is documented in the literature at 60% in CML.[84] The importance of donor T lymphocytes in graft-versus-leukemia is underscored by the ability to induce complete remissions in post-transplant relapsed CML patients with a donor lymphocyte infusion, even without the use of conditioning chemotherapy.[85–87]

It is also important to recognize that although in the US and Europe TKI therapy is considered first-line therapy, in some developing countries the cost of 1 year of imatinib outweighs the personal cost to a patient for a related-donor SCT, and for this economic reason alone, SCT still remains the first choice for many CML patients abroad.[88,89] This is an example where socioeconomic influence outweighs medical indication/advances and deserves further attention from the world medical community.

Philadelphia Chromosome Negative Myeloproliferative Neoplasms

In 2005, 4 laboratory groups nearly simultaneously published their findings on the discovery of underlying activating mutation in the Janus kinase gene, JAK2, in approximately 95%, 50%, and 50% of patients with PV, ET, and PMF, respectively, and the MPL515L/K mutation in 5% of PMF patients.[90–93] The finding of a single point mutation in a critical intracellular signaling protein that leads to constitutive cellular activity drew intense interest and provided a target for the pharmaceutical industry in designing small-molecule inhibitors that would mimic the success achieved with imatinib in CML. Janus kinase 2 belongs to a family of intracellular tyrosine kinases responsible for transducing signals via receptor activation to the nucleus. A nucleotide switch of G to T at position 1849 in exon 14 of the JAK2 pseudo-kinase domain, JH2, produces a constitutively active tyrosine kinase due to loss of the autoinhibitory control function of this domain. This leads to cytokine-independent activation and proliferation of JAK2V617F-bearing cells.[94] Mutations in exon 12 of JAK2 have also been described in JAK2V617F-negative patients with PV.[95] More than 10 deletions have been identified in exon 12, which results in a clinical phenotype of primary erythrocytosis, low serum erythropoietin, and younger age at diagnosis.[96–98] The prognosis of patients with the JAK2 exon 12 mutation is similar to that of patients bearing the JAK2V617F mutation.

Additional gene mutations that likely influence MPN phenotype have also been identified and include MPLW515L/K (5% of PMF and 1% of ET patients) and can even coexist in the same patient.[99,100] The presence of mutant ten-eleven translocation-2 (TET2) can be found in both JAK2V617F-positive and -negative MPNs (as well as other myeloid malignancies) at a frequency of approximately 12%, and does not appear to influence survival, rate of leukemic transformation, or thrombosis.[101,102] Although this gene mutation is more prevalent in older patients across each subtype of MPN, it holds limited prognostic significance. Mutations in the additional sex combs like 1 gene (ASXL1) located on the long arm of chromosome 20 are found in 10%, 10%, and 40% of patients with de novo AML, myelodysplastic syndrome, and chronic myelomonocytic leukemia, respectively, and have also been documented in cases of PMF, ET, and MF-BP.[103–105] The function of ASXL1 is poorly understood; it is believed to encode a protein responsible for the dual opposing actions of enhancing/repressing the transcription of genes including retinoic acid receptor via modulation of histone H3 methylation.[106] Larger studies are required to better understand the prognostic significance of ASXL1 gene mutations and the association of this mutation with mutations involving JAK2 and TET2. Mutations in CBL, Ezh2, and IDH have also been recently recognized in a variety of myeloid disorders including MPNs, and their pathogenetic significance is still under investigation.[107]

Janus Kinase 2 Inhibitors

Bone marrow transplant mouse models using irradiated recipient mice and murine marrow cells transduced with murine JAK2V617F-expressing retrovirus were used to investigate the role of JAK2V617F in hematopoiesis.[108,109] Initially, mice developed polycythemia and splenomegaly with bone marrow trilineage hyperplasia, and a predominance of myeloid and erythroid progenitor cells. After several months, the mice developed anemia and had evidence of marrow reticulin fibrosis. These exciting studies suggested that JAK2V617F played at least a partial role in the development of a PV and eventually an MF phenotype and provided an in vivo model for testing novel anti-JAK2 TKIs.

Close to 20 new agents are being tested in preclinical models with signs of in vitro activity. The majority of clinical research has been in patients with MF and include agents such as CEP-701 (lestaurtinib; Cephalon, Frazer, PA), INCB018424 (Incyte Corp., Wilmington, DE), TG101348 (TargeGen, San Francisco, CA), AZD1480 (AstraZeneca, Wilmington, DE), and SB1518 (S*BIO, Pte Ltd., Singapore).[110] The furthest along in development is INCB018424, which is currently being evaluated in a randomized, double-blind, placebo-controlled phase III study with primary endpoints of efficacy, safety, and tolerability and a secondary endpoint of symptom improvement as assessed by a newly developed MF-specific quality-of-life tool.[111] The agent SB1518 is a selective JAK2 and FLT-3 inhibitor that has been evaluated in a phase I dose-escalation study with clinical activity observed in some patients, and diarrhea and thrombocytopenia were noted as the most significant adverse events.[112]

> *Close to 20 Janus kinase 2 (JAK2) inhibitors are being tested in preclinical models with signs of in vitro activity. The majority of clinical research has been in patients with myelofibrosis.*

The agent CEP-701 is a nonspecific multi-TKI with activity against JAK2 and FMS-like tyrosine kinase 3 (FLT-3). It is a semisynthetic indocarbazole alkaloid derived from K-252a, a metabolite isolated from the culture broth of *Nonomurea longicatena*.[113] It was initially evaluated in the treatment of relapsing refractory FLT-3+ AML and showed evidence of activity.[114] It was found to inhibit the growth of JAK2V617F-dependent cell lines in both in vitro and xenograft models, as well as erythroid cells expanded from primary MPN CD34+ cells.[115] Inhibition of phosphorylation of STAT 3/5, AKT, and ERK and decreased expression of Bcl-xL and cyclin D1/D2 were observed and support the mechanism of JAK2-signaling inhibition.

A phase II study of CEP-701 in 22 JAK2V617F+ MF patients showed a 27% clinical improvement by International Working Group for Myelofibrosis Research and Treatment criteria.[116] However, no significant improvement was seen in bone marrow fibrosis or JAK2V617F allele burden. Anemia (14%) and thrombocytopenia (23%) were the main grade III/IV hematologic toxicities noted, and gastrointestinal side effects were the most common non-hematologic side effect seen (72% all grade and 9% grade III/IV diarrhea). In this trial, CEP-701 was used in the liquid formulation at a dose that was well below the maximally tolerated dose established in the phase I study. The MPD-RC is currently evaluating CEP-701 in a capsule formulation.[117]

Inhibitors of JAK2 have also been investigated in JAK2V617+ PV and ET. Moliterno *et al.* has reported the results of 39 high-risk JAK2V617F+ PV/ET patients treated with CEP-701.[118] Nearly 85% of patients experienced an appreciable decrease in palpable splenomegaly, and all patients had improvement in pruritus. Interestingly, 5 cases of thromboses were seen in this trial and included both arterial and venous events. This has not been observed in other malignancies, including MF treated with oral JAK2 inhibitors. At this time we can only conclude that oral JAK2 inhibitors may not be effective as a single agent alone in reducing the risk of thrombosis in PV/ET.

Not all JAK2 inhibitors have enjoyed the same level of clinical success due to concern for potential toxicity. The agent XL019 (Exelixis, San Francisco, CA) was a selective JAK2 inhibitor that appeared promising early on in the phase I study with improvements in constitutional symptoms and reduction in palpable spleen size without significant hematologic toxicity.[119] However, due to the emergence of peripheral neuropathy in several patients, this drug is not presently being pursued in active clinical trials.

Chromatin-Modifying Agents

DNA methylation and histone acetylation are normal epigenetic mechanisms that result in inheritable silencing of genes without any change in the actual gene sequence itself. Although these processes are necessary for normal mammalian cell development, if inappropriately activated they can result in the repression of tumor-suppressor gene transcription and contribute to malignant transformation.[120] The use of epigenetic directed therapy has enjoyed tremendous success in the treatment of myelodysplastic syndrome (MDS), and 2 hypomethylating agents (5-azacytidine [azacitidine; Celgene] and 5-aza-2′-deoxycytidine [decitabine; Eisai Pharmaceuticals]) are now FDA-approved in the United States for the treatment of all 5 subtypes of MDS.[121–123] Azacitidine and decitabine are DNA methyltransferase inhibitors that also have rationally based application in treatment of MPNs.

> *DNA methylation and histone acetylation are normal epigenetic mechanisms that result in inheritable silencing of genes without any change in the actual gene sequence itself. If inappropriately activated, these mechanisms can result in the repression of tumor-suppressor gene transcription and contribute to malignant transformation. The use of epigenetic directed therapy has enjoyed tremendous success in the treatment of myelodysplastic syndrome.*

The abnormal trafficking of CD34+ hematopoietic stem cells from the bone marrow of patients with

MF is well characterized and results in extramedullary hematopoiesis in the spleen and other organs.[124–126] Primary myelofibrosis CD34+ HSC infused into NOD/SCID mice preferentially home to the spleen and not the bone marrow and display reduced expression of CXC chemokine receptor CXCR4.[127] In this model, sequential treatment of PMF CD34+ HSC with 2′- deoxy-5-azacytidine followed by trichostatin A induces expression of CXCR4 and restored hematopoiesis in the bone marrow and not the spleen. These provocative findings provide scientific rationale for the use of epigenetic-directed therapy in the treatment of MPNs. In 2 phase II studies, 5-azacitidine was tested as a 7-day and 5-day course every 4 weeks in patients with refractory/relapsing MF[128,129]; although global hypomethylation of DNA was observed, no significant signal for clinical activity was appreciated. 2′-deoxy-5-azacytidine holds the potential for clinical response and has been reported to have clinical benefit in MF as well MF-BP.[130,131] Low-dose 2′-deoxy-5-azacytidine as a subcutaneous injection has been evaluated in a multicenter phase II study of 20 patients with MF.[132] A 37% response rate was seen, and 2 patients with MF-BP obtained a complete[1] and partial[1] response. A 61% mean reduction in circulating peripheral blood CD34+ cells was seen, and this was exclusive to the patients who obtained a clinical response. There was, however, no statistically significant change in CXCR4 expression. As was seen with 5-azacitidine, myelosuppression was the most common toxicity.

Vorinostat, suberoylanilide hydroxamic acid, and valproic acid are histone deacetylase (HDAC) inhibitors that have been shown to have activity in MDS/AML alone and in sequential combination with hypomethylating agents.[133,134] The agent ITF2357 (Italfarmaco, Milan, Italy) is a novel synthetic HDAC class I inhibitor that has shown activity in AML and multiple myeloma cells with minimal toxicity to normal cells.[135,136] Thrombocytopenia and gastrointestinal toxicity are the most common side effects. It has been shown to specifically down-regulate JAK2V617F protein in JAK2V617F-expressing cell lines without affecting JAK2V617F mRNA levels.[137] Downstream targets such as phosphorylated STAT-5 are also reduced, suggesting that post-transcriptional mechanisms lead to this effect. Preclinical studies utilizing purified CD34+ PMF in culture with cytokines and LBH589 (panobinostat; Novartis) resulted in both reduced number of CD34+ cells and up-regulation of CXCR4 expression in these cells.[138] In addition, a reduction in the number of JAK2V617F+ homozygous hematopoietic colonies grown in semisolid culture was seen in 2 of the 4 PMF patient samples. LBH589 is a novel potent pan-HDAC inhibitor and heat shock protein 90 inhibitor that has completed phase I testing in patients with MF and shown a signal of activity, with improvement in anemia, reduction in splenomegaly, and amelioration of constitutional symptoms.[138] Thrombocytopenia was found to be the dose-limiting toxicity, and phase II studies are now ongoing.

Interferon-alfa

The exact mechanism(s) by which IFN achieves clinical responses in patients with MPNs is not fully understood. Interferons have diverse biological functions, which include induction of pro-apoptotic genes, modulation of cellular differentiation, antiangiogenic activity, and stimulation of immunity by enhancing the cytotoxic action of various immune cells, as well as the up-regulation of major histocompatiblity complex class I antigens.[139,140] The rationale for IFN therapy in MPNs is based on the in vitro observation that IFN-alfa can markedly decrease the colony-forming ability of erythroid, granulocytic, and megakaryocytic progenitors in PV and PMF.[141–144] Documentation of conversion to polyclonal hematopoiesis with IFN treatment in PV supports the selective targeting of the malignant MPN hematopoietic clone.[145,146] In addition, IFN-alfa can directly inhibit thrombopoietin-induced megakaryopoiesis in vitro via suppression of JAK2 signaling and promotion of suppressor of cytokine signaling (SOCS).[147] Interferon had been widely used in the treatment of CML before the advent of imatinib.[148] The appreciation of the clinical efficacy of IFN therapy in PV and ET is not new and has been reported more than 20 years ago.[149,150] Recently, the pegylated form of IFN-alfa-2a has gained much attention due to the less-frequent administration and the improved side-effect profile.[151,152] In a phase II multicenter study, pegylated IFN-alfa-2a (Pegasys; Roche) in 40 PV patients induced a 95% hematologic complete remission rate, with only a small fraction of patients discontinuing treatment due to toxicity.[153,154] A molecular complete remission as defined as an undetectable JAK2V617F burden was obtained in 7 patients lasting >18+ months and was even observed to persist in patients who discontinued treatment.

Other Novel Therapeutic Approaches

B-cell lymphoma-extra large (Bcl-xL) is a Bcl-2–related protein that functions to stabilize the mitochondrial membrane potential inhibiting the release of cytochrome C and the subsequent activation of death caspases.[155] Dysregulation of the apoptotic pathway has been observed in ET CD34+- derived megakaryocytes and PV erythroid lineage cells.[156,157]

Overexpression of Bcl-xL in PV erythroid-lineage cells has been linked to impaired apoptosis in these cells.[157] Additionally, a temporal relationship between Bcl-xL expression in megakaryocytes and a change in morphology of the demarcation membrane structure is associated with proplatelet production.[158] The agent ABT-737 (Abbott Laboratories) is a small-molecule BCL-xL inhibitor and has demonstrated selective killing of JAK2V617F+ PV CD34+ cells in the presence of low-dose IFN-alfa-2a.[159] Our group has also shown that exposing PMF and PV CD34+ cells to a JAK2 inhibitor, INCB018424, in combination with ABT-737, was significantly more effective in inducing their apoptosis and reducing assayable JAK2V617F$^+$ colonies.[160] This selective inhibition of JAK2V617$^+$ MPN progenitor cells with combination treatment represents a potential therapeutic approach that may be more effective than existing single-agent JAK2 inhibitor studies. GX15-070 (obatoclax mesylate), a small-molecule pan–BCL-2 family antagonist, was investigated in a phase II study in 22 patients with MF and was associated with grade III/IV anemia (27%), thrombocytopenia (18%), and low-grade ataxia and fatigue (50%).[161] No significant clinical activity was observed at the dose and schedule evaluated in this trial. In the future, clinical trials will be designed combining agents from these various classes (JAK2 inhibitors, biologics) together to exploit different mechanisms of action (Table 1).

Table 1. *Potential Combination Therapies in Ph-Negative MPNs.*

MPN	Therapeutic Agent(s)	Class
MF	LBH589/INCB018424	HDAC inhibitor/JAK2 inhibitor
MF	LBH589/GC1008	HDAC inhibitor/ anti–TGF-β
MF	GC1008/INCB018424	Anti–TGF-β/JAK2 inhibitor
MF	LBH589/ABT-263	HDAC inhibitor/Bcl-xL inhibitor
MF	LBH589/SGI-1776	HDAC inhibitor/PIM1 inhibitor
MF	LBH589/RAD001	HDAC inhibitor/mTOR inhibitor
PV	ABT-263/IFN-alfa-2a	Bcl-xL inhibitor/biologic agent

Abbreviations: Bcl-xL, B-cell lymphoma-extra large; HDAC, histone deacetylase; IFN, interferon; JAK2, Janus kinase 2; MF, myelofibrosis; MPN, myeloproliferative neoplasm; mTOR, mammalian target of rapamycin; Ph, Philadelphia chromosome; PIM1, proto-oncogene serine/threonine-protein kinase Pim-1; PV, polycythemia vera; TGF-β, transforming growth factor β.

Transforming growth factor β (TGF-β) is a pleiotropic cytokine implicated in the promotion of angiogenesis, tumor proliferation, collagen fibrosis, metastasis, and down-regulation of antitumor immunity.[162,163] Enhanced expression of TGF-β has been seen in a variety of solid and hematologic malignancies including prostate, breast, melanoma, renal, liver, small cell lung cancer, multiple myeloma, and PMF.[164,165] In vitro studies provide evidence for the inhibitory role of TGF-β on hematopoietic colony formation by murine and human hematopoietic progenitor cells.[164] In vitro, blocking antibodies to TGF-β led to loss of quiescence by a variety of primitive hematopoietic progenitor cells including colony-forming unit (CFU) blast, CFU-mix, CFU-GM, and BFU-E and prolongation of the proliferation of long-term culture initiating cells.[164] Furthermore, TGF-β has been shown to down-regulate the expression of interleukin-1 receptor, interleukin-6 receptor, c-KIT, and c-MPL expressed by primitive human hematopoietic cells, providing a mechanism by which its inhibitory function is exerted. Recently, a critical role for the TGF-β-forkhead transcription factors (FOXO) pathway in the maintenance of leukemia-initiating cells has been reported.[166] Naka *et al.* demonstrated that FOXO3a plays a critical role in the maintenance of CML stem cells and that TGF-β controls FOXO3a nuclear localization. These data suggest that depletion of TGF-β either with a blocking antibody or with a small-molecule receptor antagonist would provide a means of eliminating cancer stem cells associated with another MPN such as MF. Clinical trials are now beginning at MSSM and Mayo Clinic Scottsdale to evaluate GC1008 (Genzyme), a human IgG Kappa 4 monoclonal antibody capable of neutralizing all TGF-β isoforms in patients with MF and ET with measurable bone marrow fibrosis.

The mammalian target of rapamycin (mTOR) is a serine/threonine kinase that is a downstream target of key signaling proteins such as Akt, PI3k, GTPase Ras, and Rheb and is constitutively activated in cells harboring the JAK2V617F mutation.[167] A specific inhibitor of the mTOR pathway, RAD001, has been shown to inhibit the proliferation of JAK2V617F-expressing cell lines and CD34+ cells from PV and PMF patients. Ongoing phase I/II studies in Italy have shown activity in improving constitutional symptoms, splenomegaly, and reduction in circulating CD34+ cells in MPN patients.[168] A serine/threonine kinase, PIM1 is responsible for promoting cell survival and proliferation at least in part through a Bcl-2–dependent mechanism.[169] Combination treatment with SGI-1776 (SuperGen Inc.), a PIM1 kinase inhibitor, and either a HDAC inhibitor (LBH589) or HSP90 inhibitor (AUY922, Novartis) has been shown in preclinical studies to result in synergistic induction of apoptosis in both MPN cell lines and primary MPN cells.[170] This represents yet another opportunity for potential combination therapy in future clinical trials where several MPN-associated activated molecular pathways are targeted to produce enhanced selective deletion of the malignant cell population.

Combination Therapy

Achieving a complete hematologic, cytogenetic, and molecular response and clinical remission requires the use of combination therapy in the treatment of the majority of hematologic malignancies, with the notable exception of single-agent imatinib in CML. It is reasonable that combination therapy will be required to achieve true success in the treatment of MPNs because the pathobiology is likely complex and not due to the activation of a single oncogenic pathway. Recently, combination therapy with the HDAC inhibitor LBH589 and a JAK2 inhibitor was shown in vitro to be even more effective than single-agent treatment in inhibiting the JAK-STAT pathway and inhibiting the proliferation of JAK2V617F-expressing cell lines and primary PMF CD34+ HSCs in vitro.[171] Based on our recent phase I results of LBH589, future proposals to explore a JAK2 inhibitor in combination with LBH589 in patients with advanced forms of MF in a phase I/II setting appear warranted. This would represent the first attempt at true combination therapy for MF in the world literature. Table 1 shows several possible drug combinations that will depend on the results from the phase I studies with ABT-263 and GC1008. This approach to the treatment of MPNs has not been utilized before, largely due to the previous limited understanding of the cellular and genetic causes of MPNs and the paucity of available relevant novel therapeutic agents. The development of combination therapies to treat MPN patients will likely be the focus of many research groups in the following years.

EXPERIMENTAL TRIALS AT MOUNT SINAI SCHOOL OF MEDICINE

In 2007, Dr Hoffman returned to MSSM to develop the MPD program, and in doing so established MSSM as the lead institution of the 38-international member MPD-RC, of which he is the principal investigator of a National Cancer Institute–sponsored program project grant. The MPD-RC consists of a basic science core, a biomarker core, and a clinical trials core. Dr Lewis Silverman is the lead investigator of the clinical trials project of the MPD-RC, and Dr John Mascarenhas is primarily responsible for the enrollment and conduct of the experimental therapeutics component of the MPD program at MSSM. There are 7 experimental treatment protocols that are available for patients with MF and 3 for patients with PV and ET. These are listed in Table 2. These are non-competing clinical trials that offer patients treatment options that are not yet available in the community. The LBH589 trial, which was described above, is an investigator-initiated phase I/II trial and is now enrolling in the phase II component of this study. The AZD1480 study is a first-in-man phase I study that is currently in the dose-escalation, dose-finding phase and is available only at MSSM, MD Anderson, and Institut Gustave Roussy in France. Enrollment has been completed for the CEP-701 trial for the PV/ET patients, and there are several patients who continue beyond a year on study with marked and continued improvements in symptoms and palpable splenomegaly. The MPD-RC CEP-701 for MF trial is now enrolling patients on the phase II portion and has shown improvements in both reduction in splenomegaly and amelioration of constitutional symptoms. The GC1008 clinical trial will open for enrollment in a phase I design in the latter part of 2010 and targets patients with measurable bone marrow fibrosis, including patients with low-risk disease. The INCB18424 study is a very large multicenter phase III study performed across several continents in patients with MF. This is the largest prospective study in MF ever conducted and is the first oral JAK2 inhibitor study to apply for FDA approval based on the pending results. The MPD-RC 101 RIC-SCT protocol is the only transplant protocol available at MSSM and has already completed accrual in the related-donor arm of the study.

CONCLUSION

The MPNs represent one of the most challenging, yet most rewarding areas of treatment within hematological malignancies. For many decades, treatment options remained limited and at best offered only a palliative approach. In recent years, we have witnessed an explosion of experimental therapeutic options for patients with MPNs. This is mostly based on the scientific advances augmenting the understanding of the molecular pathogenesis of these bone marrow malignancies. Undoubtedly, imatinib stands as a prototype of targeted therapy in CML and represents the best in translational medicine in hematologic malignancies. Now, multiple TKIs exist for the treatment of CML and offer physicians options in first-line therapy of a disease that had formerly carried a dismal prognosis. The discovery of the JAK2-related

Table 2. *Therapeutic Trials in MPNs Available at MSSM.*

Trial	Agent	Class	Phase	Sponsor	Target Population
A Phase I/II Open-Label Study of LBH589, a Novel HDACi, in Patients With Primary MF and Post-PV/ET MF	LBH589 (panobinostat)	HDAC inhibitor	II	MSSM	MF
A Phase I/II, Open-Label, Multicentre Study to Assess the Safety, Tolerability, Pharmacokinetics and Preliminary Efficacy of the JAK2 inhibitor AZD1480 Administered Orally to Patients With Primary MF and Post-PV/ET MF	AZD1480	JAK2 inhibitor	I/II	Multicenter, pharmaceutical sponsored	MF
A Multicenter, Open-Label Phase I/II Study of JAK2 Inhibitor CEP-701 (Lestaurtinib) in Adults With MF	CEP-701 (lestaurtinib)	JAK2 inhibitor	I/II	MPD-RC	MF
Phase I Study of GC1008 in Patients With Primary MF, post-PV/ET MF	GC1008	Anti–TGF-β monoclonal antibody	I	MSSM	MF
A Randomized, Double-Blind, Placebo-Controlled Phase III Study of a JAK2 Inhibitor (Incyte) Administered Orally to Subjects With Primary MF and Post-PV/ET MF	INCB18424	JAK2 inhibitor	III	Multicenter, pharmaceutical sponsored	MF
Phase III Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study to Compare Efficacy and Safety of Pomalidomide in Subjects With MPN-Associated Myelofibrosis and Red Blood Cell Transfusion Dependence	Pomalidomide	IMiD	III	Multicenter, pharmaceutical sponsored	MF
Randomized Trial of Pegylated IFN-alfa2a versus Hydroxyurea Therapy in the Treatment of High-Risk PV and ET (diagnosed <3 years)	Pegasys	Biologic	II	MPD-RC	PV/ET
Single-Arm Salvage Therapy With Pegylated IFN-alfa2a for Patients With High Risk PV or ET Who Are Either Hydroxyurea Resistant or Intolerant or Have Had Abdominal Vein Thrombosis	Pegasys	Biologic	II	MPD-RC	PV/ET
An Open-Label Study of Oral CEP-701 in Patients With PV or ET With the JAK2V617F Mutation	CEP-701 (lestautinib)	JAK2 inhibitor	II	Multicenter, pharmaceutical sponsored	PV/ET
Study of Fludarabine-Based Conditioning for Allogeneic SCT for MF	Fludarabine/melphalan ± ATG	SCT	II	MPD-RC	MF

Abbreviations: ATG, anti-thymocyte globulin; ET, essential thrombocythemia; HDAC, histone deacetylase; HDACi, histone deacetylase inhibitor; IFN, interferon; IMiD, immunomodulatory drug; JAK2, Janus kinase 2; MF, myelofibrosis; MPD-RC, Myeloproliferative Disorders Research Consortium; MPN, myeloproliferative neoplasm; MSSM, Mount Sinai School of Medicine; PV, polycythemia vera; SCT, stem cell transplantation.

mutations has resulted in the development of a number of targeted oral agents that hope to change the natural progression of the Ph-negative MPNs. As promising as the preclinical studies appeared, there is currently no evidence that any of the ongoing oral JAK2 inhibitor trials can significantly improve cytopenias, reverse marrow fibrosis, improve marrow histopathology, and change the natural clinical history of MPNs. Despite this, they may still play an important role in the armamentarium of agents used in the palliation of symptomatic splenomegaly and debilitating constitutional symptoms in patients with MPNs. In the future, innovative efforts to combine various novel therapeutic agents will likely allow for improvement in clinical responses by exploiting several active disease initiating/supporting pathways in MPNs. The search for effective agents based on rational design targeting the malignant stem cell population remains the primary goal of the translational effort in MPN research.

ACKNOWLEDGMENT

This work was supported by grants from the National Cancer Institute and the Myeloproliferative Disorders Foundation.

SUPPORTING INFORMATION

All research involving the use of human and animal subjects was conducted under the strict guidelines of the Declaration of Helsinki and approved by the Institutional Review Board/Program for the Protection of Human Subjects at Mount Sinai School of Medicine. Each subject entered on a clinical or tissue-banking study provided informed consent.

DISCLOSURES

Potential conflict of interest: Both authors declare research funding from Novartis, Cephalon, Celgene, Roche, Astra Zeneca, and Genzyme.

REFERENCES

1. Tefferi A, Vardiman JW. Classification and diagnosis of myeloproliferative neoplasms: the 2008 World Health Organization criteria and point-of-care diagnostic algorithms. *Leukemia* 2008; 22: 14–22.
2. Bennett JH. Case of hypertrophy of the spleen and liver in which death took place from suppuration of the blood. *Edinb Med Surg J* 1845; 64: 413–423.
3. Nowell PC, Hungerford DA. A minute chromosome in human chronic granulocytic leukemia. *Science* 1960; 132: 1497.
4. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* 1973; 243: 290–293.
5. Groffen J, Heisterkamp N. The BCR/ABL hybrid gene. *Baillieres Clin Haematol* 1987; 1: 983–999.
6. Ben-Neriah Y, Daley GQ, Mes-Masson AM, et al. The chronic myelogenous leukemia-specific P210 protein is the product of the bcr/abl hybrid gene. *Science* 1986; 233: 212–214.
7. Gale RP, Cannani E. The molecular biology of chronic myelogenous leukaemia. *Br J Haematol* 1985; 60: 395–408.
8. Vaquez H. [Sur une forme spéciale de cyanose s'accompagnant d'hyperglobulie excessive et persistante]. *CR Soc Biol (Paris)* 1892; 44: 384–338. French.
9. Osler W. Chronic cyanosis with polycythaemia and enlarged spleen: a new clinical entity. *Am J Med Sci* 1903; 1: 143–146.
10. Dameshek W. Some speculations on the myeloproliferative syndromes. *Blood* 1951; 6: 372–375.
11. Wasserman LR. Polycythemia Vera Study Group: a historical perspective. *Semin Hematol* 1986; 23: 183–187.
12. Berk PD, Goldberg JD, Donovan PB, et al. Therapeutic recommendations in polycythemia vera based on Polycythemia Vera Study Group protocols. *Semin Hematol* 1986; 23: 132–143.
13. Fruchtman SM, Mack K, Kaplan ME, et al. From efficacy to safety: a Polycythemia Vera Study group report on hydroxyurea in patients with polycythemia vera. *Semin Hematol* 1997; 34: 17–23.
14. Barosi G, Mesa RA, Thiele J, et al. Proposed criteria for the diagnosis of post-polycythemia vera and post-essential thrombocythemia myelofibrosis: a consensus statement from the International Working Group for Myelofibrosis Research and Treatment. *Leukemia* 2008; 22: 437–438.
15. Swerdlow SH, Campo E, Harris NL, et al., eds. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. 4th ed. Geneva, Switzerland: World Health Organization Press; 2008.
16. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. *CA Cancer J Clin* 2006; 56: 106–130.
17. Sawyers CL. Chronic myeloid leukemia. *N Engl J Med* 1999; 340: 1330–1340.
18. Savage DG, Szydlo RM, Goldman JM. Clinical features at diagnosis in 430 patients with chronic myeloid leukaemia seen at a referral centre over a 16-year period. *Br J Haematol* 1997; 96: 111–116.
19. Kantarjian H, Sawyers C, Hochhaus A, et al. Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia. *N Engl J Med* 2002; 346: 645–652.
20. Faderl S, Kantarjian HM, Talpaz M. Chronic myelogenous leukemia: update on biology and treatment. *Oncology (Williston Park)* 1999; 13: 169–184.
21. Interferon alfa versus chemotherapy for chronic myeloid leukemia: a meta-analysis of seven randomized trials. Chronic Myeloid Leukemia Trialists' Collaborative Group. *J Natl Cancer Inst* 1997; 89: 1616–1620.

22. Gangat N, Wolanskyj AP, McClure RF, et al. Risk stratification for survival and leukemic transformation in essential thrombocythemia: a single institutional study of 605 patients. *Leukemia* 2007; 21: 270–276.
23. Rozman C, Giralt M, Feliu E, et al. Life expectancy of patients with chronic nonleukemic myeloproliferative disorders. *Cancer* 1991; 67: 2658–2663.
24. Passamonti F, Rumi E, Caramella M, et al. A dynamic prognostic model to predict survival in post-polycythemia vera myelofibrosis. *Blood* 2008; 111: 3383–3387.
25. Cervantes F, Passamonti F, Barosi G. Life expectancy and prognostic factors in the classic BCR/ABL-negative myeloproliferative disorders. *Leukemia* 2008; 22: 905–914.
26. Tefferi A. Risk-based management in essential thrombocythemia. In: *Hematology 1999*. Washington, DC: American Society of Hematology; 1999; 172–177.
27. Dupriez B, Morel P, Demory JL, et al. Prognostic factors in agnogenic myeloid metaplasia: a report on 195 cases with a new scoring system. *Blood* 1996; 88: 1013–1018.
28. Cervantes F, Dupriez B, Pereira A, et al. New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment. *Blood* 2009; 113: 2895–2901.
29. Mascarenhas J, Hoffman R. Risk adapted approach to the treatment of primary myelofibrosis. Hematology Education: the education program for the annual congress of the European Hematology Association. 2009; 3: 192–199.
30. Tefferi A, Cortes J, Verstovsek S, et al. Lenalidomide therapy in myelofibrosis with myeloid metaplasia. *Blood* 2006; 108: 1158–1164.
31. Barosi G, Elliott M, Canepa L, et al. Thalidomide in myelofibrosis with myeloid metaplasia: a pooled-analysis of individual patient data from five studies. *Leuk Lymphoma* 2002; 43: 2301–2307.
32. Mesa RA, Steensma DP, Pardanani A, et al. A phase 2 trial of combination low-dose thalidomide and prednisone for the treatment of myelofibrosis with myeloid metaplasia. *Blood* 2003; 101: 2534–2541.
33. Mesa R, Yao X, Cripe LD, et al. Lenalidomide and prednisone for primary and post polycythemia vera/essential thrombocythemia myelofibrosis (MF): an Eastern Cooperative Oncology Group (ECOG) phase II trial. Paper presented at: American Society of Hematology Annual Meeting; December 6, 2008;; San Francisco, CA.
34. Cervantes F, Alvarez-Larrán A, Domingo A, et al. Efficacy and tolerability of danazol as a treatment for the anaemia of myelofibrosis with myeloid metaplasia: long-term results in 30 patients. *Br J Haematol* 2005; 129: 771–775.
35. Cervantes F, Alvarez-Larrán A, Hernández-Boluda JC, et al. Erythropoietin treatment of the anaemia of myelofibrosis with myeloid metaplasia: results in 20 patients and review of the literature. *Br J Haematol* 2004; 127: 399–403.
36. Rodríguez JN, Martino ML, Diéguez JC, et al. rHuEpo for the treatment of anemia in myelofibrosis with myeloid metaplasia: experience in 6 patients and meta-analytical approach. *Haematologica* 1998; 83: 616–621.
37. Barosi G, Liberato LN, Guarnone R. Serum erythropoietin in patients with myelofibrosis with myeloid metaplasia. *Br J Haematol* 1993; 83: 365–369.
38. Hasselbalch HC, Clausen NT, Jensen BA. Successful treatment of anemia in idiopathic myelofibrosis with recombinant human erythropoietin. *Am J Hematol* 2002; 70: 92–99.
39. Lofvenberg E, Wahlin A. Management of polycythaemia vera, essential thrombocythaemia and myelofibrosis with hydroxyurea. *Eur J Haematol* 1988; 41: 375–381.
40. Manoharan A, Pitney WR. Chemotherapy resolves symptoms and reverses marrow fibrosis in myelofibrosis. *Scand J Haematol* 1984; 33: 453–459.
41. Petti MC, Latagliata R, Spadea T, et al. Melphalan treatment in patients with myelofibrosis with myeloid metaplasia. *Br J Haematol* 2002; 116: 576–581.
42. Tefferi A, Silverstein MN, Li CY. 2-Chlorodeoxyadenosine treatment after splenectomy in patients who have myelofibrosis with myeloid metaplasia. *Br J Haematol* 1997; 99: 352–357.
43. Rondelli D, Barosi G, Bacigalupo A, et al. Allogeneic hematopoietic stem-cell transplantation with reduced-intensity conditioning in intermediate- or high-risk patients with myelofibrosis with myeloid metaplasia. *Blood* 2005; 105: 4115–4119.
44. Guardiola P, Anderson JE, Bandini G, et al. Allogeneic stem cell transplantation for agnogenic myeloid metaplasia: a European Group for Blood and Marrow Transplantation, Société Française de Greffe de Moelle, Gruppo Italiano per il Trapianto del Midollo Osseo, and Fred Hutchinson Cancer Research Center Collaborative Study. *Blood* 1999; 93: 2831–2838.
45. Rondelli D, Boyer MW, Isola LM, et al. First prospective study of allogeneic peripheral blood stem cell (PBSC) transplantation in patients with myelofibrosis in the United States: interim analysis of MPD-RC 101 protocol. *J Clin Oncol* 2010; 28(suppl 15S): Abstract 6535.
46. O'Brien SG, Guilhot F, Larson RA, et al. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med* 2003; 348: 994–1004.
47. Deininger MW, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. *Blood* 2000; 96: 3343–3356.
48. Deininger M, O'Brien SG, Guilhot F, et al. International randomized study of interferon vs STI571 (IRIS) 8-year follow up: Sustained survival and low risk for progression or events in patients with newly diagnosed chronic myeloid leukemia in chronic phase (CML-CP) treated with imatinib. Paper presented at: American Society of Hematology Annual Meeting; December 5–8, 2009; San Francisco, CA. Abstract 1126.
49. Cortes JE, Jones D, O'Brien S, et al. Results of dasatinib therapy in patients with early chronic-phase chronic myeloid leukemia. *J Clin Oncol* 2010; 28: 398–404.
50. Cortes JE, Jones D, O'Brien S, et al. Nilotinib as front-line treatment for patients with chronic myeloid leukemia in early chronic phase. *J Clin Oncol* 2010; 28: 392–397.
51. Saglio G, Kim DW, Issaragrisil S, et al. Nilotinib versus imatinib for newly diagnosed chronic myeloid leukemia. *N Engl J Med* 2010; 362: 2251–2259.

52. Kantarjian H, Shah NP, Hochhaus A, et al. Dasatinib versus imatinib in newly diagnosed chronic-phase chronic myeloid leukemia. *N Engl J Med* 2010; 362: 2260–2270.
53. Quintás-Cardama A, Kantarjian HM, Jones D, et al. Delayed achievement of molecular responses is associated with increased risk of progression among patients with chronic myelogenous leukemia in chronic phase treated with imatinib. *Blood* 2009; 113: 6315–6321.
54. Shah NP, Nicoll JM, Nagar B, et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2002; 2: 117–125.
55. Gorre ME, Mohammed M, Ellwood K, et al. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. *Science* 2001; 293: 876–880.
56. Apperley JF. Part I: mechanisms of resistance to imatinib in chronic myeloid leukaemia. *Lancet Oncol* 2007; 8: 1018–1029.
57. O'Hare T, Walters DK, Stoffregen EP, et al. In vitro activity of Bcr-Abl inhibitors AMN107 and BMS-354825 against clinically relevant imatinib-resistant Abl kinase domain mutants. *Cancer Res* 2005; 65: 4500–4505.
58. Wang L, Giannoudis A, Lane S, et al. Expression of the uptake drug transporter hOCT1 is an important clinical determinant of the response to imatinib in chronic myeloid leukemia. *Clin Pharmacol Ther* 2008; 83: 258–264.
59. White DL, Saunders VA, Dang P, et al. Most CML patients who have a suboptimal response to imatinib have low OCT-1 activity: higher doses of imatinib may overcome the negative impact of low OCT-1 activity. *Blood* 2007; 110: 4064–4072.
60. Cortes JE, Kantarjian H, Brümmendorf T, et al. Safety and efficacy of bosutinib (SKI-606) in patients with chronic phase chronic myeloid leukemia following resistance or intolerance to imatinib. *J Clin Oncol* 2010; 28(suppl 15S): Abstract 6502.
61. Cortes J, Paquette R, Talpaz M, et al. Preliminary clinical activity in a phase I trial of the BCR-ABL/IGF-1R/Aurora kinase inhibitor XL228 in patients with Ph^{+}+ leukemias with either failure to multiple TKI therapies or with T315I mutation. *Blood* 2008; 112: Abstract 3232.
62. Rousselot P, Huguet F, Rea D, et al. Imatinib mesylate discontinuation in patients with chronic myelogenous leukemia in complete molecular remission for more than 2 years. *Blood* 2007; 109: 58–60.
63. Gontarewicz A, Balabanov S, Keller G, et al. Simultaneous targeting of Aurora kinases and Bcr-Abl kinase by the small molecule inhibitor PHA-739358 is effective against imatinib-resistant BCR-ABL mutations including T315I. *Blood* 2008; 111: 4355–4364.
64. Copland M, Pellicano F, Richmond L, et al. BMS-214662 potently induces apoptosis of chronic myeloid leukemia stem and progenitor cells and synergizes with tyrosine kinase inhibitors. *Blood* 2008; 111: 2843–2853.
65. Jørgensen HG, Holyoake TL. Characterization of cancer stem cells in chronic myeloid leukaemia. *Biochem Soc Trans* 2007; 35(part 5): 1347–1351.
66. Chen Y, Hu Y, Michaels S, et al. Inhibitory effects of omacetaxine on leukemic stem cells and BCR-ABL-induced chronic myeloid leukemia and acute lymphoblastic leukemia in mice. *Leukemia* 2009; 23: 1446–1454.
67. Cathcart K, Pinilla-Ibarz J, Korontsvit T, et al. A multivalent bcr-abl fusion peptide vaccination trial in patients with chronic myeloid leukemia. *Blood* 2004; 103: 1037–1042.
68. Abruzzese E, Trawinska MM, Coletta A, et al. Peptide-vaccine treatment associated with TKI therapy in patients with CML is able to induce immunologic, cytogenetic and molecular responses: a single center experience with long-term follow up. *Blood* 2009; 114: Abstract 2185.
69. Bocchia M, Gentili S, Abruzzese E, et al. Effect of a p210 multipeptide vaccine associated with imatinib or interferon in patients with chronic myeloid leukaemia and persistent residual disease: a multicentre observational trial. *Lancet* 2005; 365: 657–662.
70. Li Z, Qiao Y, Liu B, et al. Combination of imatinib mesylate with autologous leukocyte-derived heat shock protein and chronic myelogenous leukemia. *Clin Cancer Res* 2005; 11: 4460–4468.
71. Marin D, Mauro M, Goldman J, et al. Preliminary results from a phase 2 trial of AG-858, an autologous heat shock protein-peptide vaccine, in combination with imatinib in patients with chronic phase chronic myeloid leukemia (CML) resistant to prior imatinib monotherapy. *Blood* 2005; 106: Abstract 1094.
72. Gale RP, Hehlmann R, Zhang MJ, et al. Survival with bone marrow transplantation versus hydroxyurea or interferon for chronic myelogenous leukemia. The German CML Study Group. *Blood* 1998; 91: 1810–1819.
73. Clift RA, Storb R. Marrow transplantation for CML: the Seattle experience. *Bone Marrow Transplant* 1996; 17(suppl 3): S1–S3.
74. Horowitz MM, Rowlings PA, Passweg JR. Allogeneic bone marrow transplantation for CML: a report from the International Bone Marrow Transplant Registry. *Bone Marrow Transplant* 1996; 17(suppl 3): S5–S6.
75. Weisdorf DJ, Anasetti C, Antin JH, et al. Allogeneic bone marrow transplantation for chronic myelogenous leukemia: comparative analysis of unrelated versus matched sibling donor transplantation. *Blood* 2002; 99: 1971–1977.
76. Goldman JM, Apperley JF, Jones L, et al. Bone marrow transplantation for patients with chronic myeloid leukemia. *N Engl J Med* 1986; 314: 202–207.
77. Hehlmann R, Berger U, Pfirrmann M, et al. Drug treatment is superior to allografting as first-line therapy in chronic myeloid leukemia. *Blood* 2007; 109: 4686–4692.
78. Bittencourt H, Funke V, Fogliatto L, et al. Imatinib mesylate versus allogeneic BMT for patients with chronic myeloid leukemia in first chronic phase. *Bone Marrow Transplant* 2008; 42: 597–600.
79. Basak G, Torosian T, Snarski E, et al. Hematopoietic stem cell transplantation for T315I-mutated chronic myelogenous leukemia. *Ann Transplant* 2010; 15: 68–70.
80. Gratwohl A, Baldomero H, Horisberger B, et al. Current trends in hematopoietic stem cell transplantation in Europe. *Blood* 2002; 100: 2374–2386.

81. Gratwohl A, Brand R, Apperley J, et al. Allogeneic hematopoietic stem cell transplantation for chronic myeloid leukemia in Europe 2006: transplant activity, long-term data and current results. An analysis by the Chronic Leukemia Working Party of the European Group for Blood and Marrow Transplantation (EBMT). *Haematologica* 2006; 91: 513–521.
82. Radich JP, Gooley T, Bensinger W, et al. HLA-matched related hematopoietic cell transplantation for chronic-phase CML using a targeted busulfan and cyclophosphamide preparative regimen. *Blood* 2003; 102: 31–35.
83. Van Rhee F, Szydlo RM, Hermans J, et al. Long-term results after allogeneic bone marrow transplantation for chronic myelogenous leukemia in chronic phase: a report from the Chronic Leukemia Working Party of the European Group for Blood and Marrow Transplantation. *Bone Marrow Transplant* 1997; 20: 553–560.
84. Goldman JM, Gale RP, Horowitz MM, et al. Bone marrow transplantation for chronic myelogenous leukemia in chronic phase: increased risk for relapse associated with T-cell depletion. *Ann Intern Med* 1988; 108: 806–814.
85. Kolb HJ, Mittermüller J, Clemm C, et al. Donor leukocyte transfusions for treatment of recurrent chronic myelogenous leukemia in marrow transplant patients. *Blood* 1990; 76: 2462–2465.
86. Drobyski WR, Keever CA, Roth MS, et al. Salvage immunotherapy using donor leukocyte infusions as treatment for relapsed chronic myelogenous leukemia after allogeneic bone marrow transplantation: efficacy and toxicity of a defined T-cell dose. *Blood* 1993; 82: 2310–2318.
87. Porter DL, Roth MS, McGarigle C, et al. Induction of graft-versus-host disease as immunotherapy for relapsed chronic myeloid leukemia. *N Engl J Med* 1994; 330: 100–106.
88. Ruiz-Argüelles GJ, Tarin-Arzaga LC, Gonzalez-Carrillo ML, et al. Therapeutic choices in patients with Ph-positive CML living in Mexico in the tyrosine kinase inhibitor era: SCT or TKIs? *Bone Marrow Transplant* 2008; 42: 23–28.
89. Gajewski JL, Robinson P. Do affluent societies have the only options for the best therapy? *Leukemia* 2007; 21: 387–388.
90. James C, Ugo V, Le Couedic JP, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* 2005; 434: 1144–1148.
91. Baxter EJ, Scott LM, Campbell PJ, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* 2005; 365: 1054–1061.
92. Kralovics R, Passamonti F, Buser AS, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med* 2005; 352: 1779–1790.
93. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell* 2005; 7: 387–397.
94. Kundrapu K, Colenberg L, Duhé RJ. Activation loop tyrosines allow the JAK2(V617F) mutant to attain hyperactivation. *Cell Biochem Biophys* 2008; 52: 103–112.
95. Scott LM, Tong W, Levine RL, et al. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *N Engl J Med* 2007; 356: 459–468.
96. Pardanani A, Lasho TL, Finke C, et al. Prevalence and clinicopathologic correlates of JAK2 exon 12 mutations in JAK2V617F-negative polycythemia vera. *Leukemia* 2007; 21: 1960–1963.
97. Pietra D, Li S, Brisci A, et al. Somatic mutations of JAK2 exon 12 in patients with JAK2 (V617F)-negative myeloproliferative disorders. *Blood* 2008; 111: 1686–1689.
98. Butcher CM, Hahn U, To LB, et al. Two novel JAK2 exon 12 mutations in JAK2V617F-negative polycythaemia vera patients. *Leukemia* 2008; 22: 870–873.
99. Lasho TL, Pardanani A, McClure RF, et al. Concurrent MPL515 and JAK2V617F mutations in myelofibrosis: chronology of clonal emergence and changes in mutant allele burden over time. *Br J Haematol* 2006; 135: 683–687.
100. Pardanani AD, Levine RL, Lasho T, et al. MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. *Blood* 2006; 108: 3472–3476.
101. Delhommeau F, Dupont S, Della Valle V, et al. Mutation in TET2 in myeloid cancers. *N Engl J Med* 2009; 360: 2289–2301.
102. Tefferi A, Pardanani A, Lim KH, et al. TET2 mutations and their clinical correlates in polycythemia vera, essential thrombocythemia and myelofibrosis. *Leukemia* 2009; 23: 905–911.
103. Carbuccia N, Murati A, Trouplin V, et al. Mutations of ASXL1 gene in myeloproliferative neoplasms. *Leukemia* 2009; 23: 2183–2186.
104. Gelsi-Boyer V, Trouplin V, Adelaide J, et al. Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia. *Br J Haematol* 2009; 145: 788–800.
105. Boultwood J, Perry J, Pellagatti A, et al. Frequent mutation of the polycomb-associated gene ASXL1 in the myelodysplastic syndromes and in acute myeloid leukemia. *Leukemia* 2010; 24: 1062–1065.
106. Lee SW, Cho YS, Na JM, et al. ASXL1 represses retinoic acid receptor-mediated transcription through associating with HP1 and LSD1. *J Biol Chem* 2010; 285: 18–29.
107. Tefferi A. Novel mutations and their functional and clinical relevance in myeloproliferative neoplasms: JAK2, MPL, TET2, ASXL1, CBL, IDH and IKZF1. *Leukemia* 2010; 24: 1128–1138.
108. Lacout C, Pisani DF, Tulliez M, et al. JAK2V617F expression in murine hematopoietic cells leads to MPD mimicking human PV with secondary myelofibrosis. *Blood* 2006; 108: 1652–1660.
109. Wernig G, Mercher T, Okabe R, et al. Expression of Jak2V617F causes a polycythemia vera–like disease with associated myelofibrosis in a murine bone marrow transplant model. *Blood* 2006; 107: 4274–4281.
110. Mesa RA. New drugs for the treatment of myelofibrosis. *Curr Hematol Malig Rep* 2010; 5: 15–21.
111. Mesa RA, Kantarjian H, Tefferi A, et al. Validation of the serial use of the Myelofibrosis Symptom Assessment Form (MF-SAF) for measuring symptomatic improvement: performance in 86 myelofibrosis

patients on INCB018424 clinical trial. *Blood* 2009; 114: Abstract 3917.

112. Verstovsek S, Odenike O, Scott B, et al. Phase I dose-escalation trial of SB1518, a novel JAK2/FLT3 inhibitor, in acute and chronic myeloid diseases, including primary or post–essential thrombocythemia/polycythemia vera myelofibrosis. *Blood* 2009; 114: Abstract 3905.
113. Kase H, Iwahashi K, Matsuda Y. K-252a, a potent inhibitor of protein kinase C from microbial origin. *J Antibiot (Tokyo)* 1986; 39: 1059–1065.
114. Smith BD, Levis M, Beran M, et al. Single-agent CEP-701, a novel FLT3 inhibitor, shows biologic and clinical activity in patients with relapsed or refractory acute myeloid leukemia. *Blood* 2004; 103: 3669–3676.
115. Hexner EO, Serdikoff C, Jan M, et al. Lestaurtinib (CEP701) is a JAK2 inhibitor that suppresses JAK2/STAT5 signaling and the proliferation of primary erythroid cells from patients with myeloproliferative disorders. *Blood* 2008; 111: 5663–5671.
116. Santos FP, Kantarjian HM, Jain N, et al. Phase 2 study of CEP-701, an orally available JAK2 inhibitor, in patients with primary or post–polycythemia vera/essential thrombocythemia myelofibrosis. *Blood* 2010; 115: 1131–1136.
117. Hexner E, Goldberg JD, Prchal JT, et al. A multicenter, open label phase I/II study of CEP701 (lestaurtinib) in adults with myelofibrosis; a report on phase I: a study of the Myeloproliferative Disorders Research Consortium (MPD-RC). Paper presented at: American Society of Hematology Annual Meeting; December 7, 2009; New Orleans, LA.
118. Moliterno AR, Hexner E, Roboz GJ, et al. An open-label study of CEP-701 in patients with JAK2 V617F-positive PV and ET: update of 39 enrolled patients. *Blood* 2009; 114: Abstract 753.
119. Shah NP, Olszynski P, Sokol L, et al. A phase I study of XL019, a selective JAK2 inhibitor, in patients with primary myelofibrosis, post–polycythemia vera, or post–essential thrombocythemia myelofibrosis. *Blood* 2008; 112: Abstract 98.
120. Esteller M. Epigenetics in cancer. *N Engl J Med* 2008; 358: 1148–1159.
121. Silverman LR, McKenzie DR, Peterson BL, et al. Further analysis of trials with azacitidine in patients with myelodysplastic syndrome: studies 8421, 8921, and 9221 by the Cancer and Leukemia Group B. *J Clin Oncol* 2006; 24: 3895–3903.
122. Kantarjian H, Issa JP, Rosenfeld CS, et al. Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. *Cancer* 2006; 106: 1794–1803.
123. Silverman LR, Holland JF, Weinberg RS, et al. Effects of treatment with 5-azacytidine on the in vivo and in vitro hematopoiesis in patients with myelodysplastic syndromes. *Leukemia* 1993; 7(suppl 1): 21–29.
124. Xu M, Bruno E, Chao J, et al. The constitutive mobilization of bone marrow–repopulating cells into the peripheral blood in idiopathic myelofibrosis. *Blood* 2005; 105: 1699–1705.
125. Barosi G, Viarengo G, Pecci A, et al. Diagnostic and clinical relevance of the number of circulating CD34(+) cells in myelofibrosis with myeloid metaplasia. *Blood* 2001; 98: 3249–3255.
126. Xu M, Bruno E, Chao J, et al. Constitutive mobilization of CD34+ cells into the peripheral blood in idiopathic myelofibrosis may be due to the action of a number of proteases. *Blood* 2005; 105: 4508–4515.
127. Wang X, Zhang W, Ishii T, et al. Correction of the abnormal trafficking of primary myelofibrosis CD34+ cells by treatment with chromatin-modifying agents. *Cancer Res* 2009; 69: 7612–7618.
128. Quintás-Cardama A, Tong W, Kantarjian H, et al. A phase II study of 5-azacitidine for patients with primary and post–essential thrombocythemia/polycythemia vera myelofibrosis. *Leukemia* 2008; 22: 965–970.
129. Mesa RA, Verstovsek S, Rivera C, et al. 5-Azacitidine has limited therapeutic activity in myelofibrosis. *Leukemia* 2009; 23: 180–182.
130. Danilov AV, Relias V, Feeney DM, et al. Decitabine is an effective treatment of idiopathic myelofibrosis. *Br J Haematol* 2009; 145: 131–132.
131. Mascarenhas J, Navada S, Malone A, et al. Therapeutic options for patients with myelofibrosis in blast phase. *Leuk Res* 2010; 34: 1246–1249.
132. Odenike OM, Godwin JE, Van Besien K, et al. Phase II trial of low dose, subcutaneous decitabine in myelofibrosis. *Blood* 2008; 112: Abstract 2809.
133. Silverman LR, Verma A, Odchimar-Reissig R, et al. A phase I/II study of vorinostat, an oral histone deacetylase inhibitor, in combination with azacitidine in patients with the myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML). Initial results of the phase I trial: A New York Cancer Consortium. *J Clin Oncol* 2008; 26(suppl): Abstract 7000.
134. Garcia-Manero G, Yang H, Bueso-Ramos C, et al. Phase 1 study of the histone deacetylase inhibitor vorinostat (suberoylanilide hydroxamic acid [SAHA]) in patients with advanced leukemias and myelodysplastic syndromes. *Blood* 2008; 111: 1060–1066.
135. Galli M, Salmoiraghi S, Golay J, et al. A phase II multiple dose clinical trial of histone deacetylase inhibitor ITF2357 in patients with relapsed or progressive multiple myeloma: preliminary results. *Blood* 2007; 110: Abstract 1175.
136. Golay J, Cuppini L, Leoni F, et al. The histone deacetylase inhibitor ITF2357 has anti-leukemic activity in vitro and in vivo and inhibits IL-6 and VEGF production by stromal cells. *Leukemia* 2007; 21: 1892–1900.
137. Guerini V, Barbui V, Spinelli O, et al. Selective targeting of the JAK2V617F mutation in polycythemia vera and essential thrombocythemia by ITF2357, a novel histone deacetylase inhibitor. *Blood* 2007; 110: Abstract 555.
138. Mascarenhas J, Wang X, Rodriguez A, et al. A phase I study of LBH589, a novel histone deacetylase inhibitor in patients with primary myelofibrosis (PMF) and post–polycythemia/essential thrombocythemia myelofibrosis (Post-PV/ET MF). *Blood* 2009; 114: Abstract 308.
139. Pestka S, Krause CD, Walter MR. Interferons, interferon-like cytokines, and their receptors. *Immunol Rev* 2004; 202: 8–32.
140. Ortaldo JR, Mason A, Rehberg E, et al. Effects of recombinant and hybrid recombinant human leukocyte interferons on cytotoxic activity of natural killer cells. *J Biol Chem* 1983; 258: 15011–15015.
141. Carlo-Stella C, Cazzola M, Gasner A, et al. Effects of recombinant alpha and gamma interferons on the in vitro growth of circulating hematopoietic

progenitor cells (CFU-GEMM, CFU-Mk, BFU-E, and CFU-GM) from patients with myelofibrosis with myeloid metaplasia. *Blood* 1987; 70: 1014–1019.

142. Castello G, Lerza R, Cerruti A, et al. The in vitro and in vivo effect of recombinant interferon alpha-2a on circulating haemopoietic progenitors in polycythaemia vera. *Br J Haematol* 1994; 87: 621–623.
143. Dudley JM, Westwood N, Leonard S, et al. Primary polycythaemia: positive diagnosis using the differential response of primitive and mature erythroid progenitors to erythropoietin, interleukin 3 and alpha-interferon. *Br J Haematol* 1990; 75: 188–194.
144. Hino M, Futami E, Okuno S, et al. Possible selective effects of interferon alpha-2b on a malignant clone in a case of polycythemia vera. *Ann Hematol* 1993; 66: 161–162.
145. Massaro P, Foa P, Pomati M, et al. Polycythemia vera treated with recombinant interferon-alpha 2a: evidence of a selective effect on the malignant clone. *Am J Hematol* 1997; 56: 126–128.
146. Liu E, Jelinek J, Pastore YD, et al. Discrimination of polycythemias and thrombocytoses by novel, simple, accurate clonality assays and comparison with PRV-1 expression and BFU-E response to erythropoietin. *Blood* 2003; 101: 3294–3301.
147. Wang Q, Miyakawa Y, Fox N, et al. Interferon-alpha directly represses megakaryopoiesis by inhibiting thrombopoietin-induced signaling through induction of SOCS-1. *Blood* 2000; 96: 2093–2099.
148. Talpaz M, Kantarjian HM, McCredie K, et al. Hematologic remission and cytogenetic improvement induced by recombinant human interferon alpha A in chronic myelogenous leukemia. *N Engl J Med* 1986; 314: 1065–1069.
149. Bellucci S, Harousseau JL, Brice P, et al. Treatment of essential thrombocythaemia by alpha 2a interferon. *Lancet* 1988; 2: 960–961.
150. Silver RT. Recombinant interferon-alpha for treatment of polycythaemia vera. *Lancet* 1988; 2: 403.
151. Glue P, Fang JW, Rouzier-Panis R, et al. Pegylated interferon-alpha2b: pharmacokinetics, pharmacodynamics, safety, and preliminary efficacy data. Hepatitis C Intervention Therapy Group. *Clin Pharmacol Ther* 2000; 68: 556–567.
152. Michallet M, Maloisel F, Delain M, et al. Pegylated recombinant interferon alpha-2b vs recombinant interferon alpha-2b for the initial treatment of chronic-phase chronic myelogenous leukemia: a phase III study. *Leukemia* 2004; 18: 309–315.
153. Kiladjian JJ, Chomienne C, Fenaux P. Interferon-alpha therapy in bcr-abl–negative myeloproliferative neoplasms. *Leukemia* 2008; 22: 1990–1998.
154. Kiladjian JJ, Cassinat B, Chevret S, et al. Pegylated interferon-alfa-2a induces complete hematologic and molecular responses with low toxicity in polycythemia vera. *Blood* 2008; 112: 3065–3072.
155. Boise LH, González-Garcia M, Postema CE, et al. Bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell* 1993; 74: 597–608.
156. Tenedini E, Fagioli ME, Vianelli N, et al. Gene expression profiling of normal and malignant CD34-derived megakaryocytic cells. *Blood* 2004; 104: 3126–3135.
157. Silva M, Richard C, Benito A, et al. Expression of Bcl-x in erythroid precursors from patients with polycythemia vera. *N Engl J Med* 1998; 338: 564–571.
158. Kaluzhny Y, Yu G, Sun S, et al. BclxL overexpression in megakaryocytes leads to impaired platelet fragmentation. *Blood* 2002; 100: 1670–1678.
159. Lu M, Zhang W, Yoo D, et al. Treatment with pegylated interferon alpha 2a in combination with the Bcl-Xl inhibitor ABT-737 specifically targets JAK2V617F positive hematopoietic progenitor cells from patients with polycythemia vera. *Blood* 2009; 114: Abstract 3916.
160. Lu M, Zhang W, Wang J, et al. Treatment in vitro with a combination of Bcl-Xl inhibitor-ABT-737 and a JAK2 inhibitor selectively eliminates JAK2V617F MPN progenitor cells. *Blood* 2009; 114: Abstract 752.
161. Parikh S, Kantarjian H, Schimmer A, et al. Phase II study of obatoclax mesylate (GX15-070), a small-molecule BCL-2 family antagonist, for patients with myelofibrosis. *Clin Lymphoma Myeloma Leuk* 2010; 10: 285–289.
162. Blobe GC, Schiemann WP, Lodish HF. Role of transforming growth factor beta in human disease. *N Engl J Med* 2000; 342: 1350–1358.
163. Dumont N, Arteaga CL. Targeting the TGF beta signaling network in human neoplasia. *Cancer Cell* 2003; 3: 531–536.
164. Fortunel NO, Hatzfeld A, Hatzfeld JA. Transforming growth factor-beta: pleiotropic role in the regulation of hematopoiesis. *Blood* 2000; 96: 2022–2036.
165. Martyre MC, Romquin N, Le Bousse-Kerdiles MC, et al. Transforming growth factor-beta and megakaryocytes in the pathogenesis of idiopathic myelofibrosis. *Br J Haematol* 1994; 88: 9–16.
166. Naka K, Hoshii T, Muraguchi T, et al. TGF-beta-FOXO signalling maintains leukaemia-initiating cells in chronic myeloid leukaemia. *Nature* 2010; 463: 676–680.
167. Vannucchi AM, Bogani C, Bartalucci N, et al. The mTOR inhibitor, RAD001, inhibits the growth of cells from patients with myeloproliferative neoplasms. *Blood* 2009; 114: Abstract 2914.
168. Vannucchi AM, Guglielmelli P, Gattoni E, et al. RAD001, an inhibitor of mTOR, shows clinical activity in a phase I/II study in patients with primary myelofibrosis (PMF) and post polycythemia vera/essential thrombocythemia myelofibrosis (PPV/PET MF). *Blood* 2009; 114: Abstract 307.
169. Lilly M, Sandholm J, Cooper JJ, et al. The PIM-1 serine kinase prolongs survival and inhibits apoptosis-related mitochondrial dysfunction in part through a bcl-2-dependent pathway. *Oncogene* 1999; 18: 4022–4031.
170. Fiskus WC, Buckley KM, Rao R, et al. Synergistic activity of co-treatment with PIM1 kinase inhibitor SGI-1776 and histone deacetylase inhibitor panobinostat or heat shock protein 90 inhibitor AUY922 against human CML and myeloproliferative neoplasm (MPN) cells. *Blood* 2009; 114: Abstract 2651.
171. Wang Y, Fiskus W, Chong DG, et al. Cotreatment with panobinostat and JAK2 inhibitor TG101209 attenuates JAK2V617F levels and signaling and exerts synergistic cytotoxic effects against human myeloproliferative neoplastic cells. *Blood* 2009; 114: 5024–5033.

Head and Neck Squamous Cell Carcinoma: New Translational Therapies

Anthony Prince, MD, Julio Aguirre-Ghizo, PhD, Eric Genden, MD, Marshall Posner, MD, and Andrew Sikora, MD, PhD

Mount Sinai School of Medicine, New York, NY

OUTLINE

Biology and Epidemiology of Head and Neck Cancer
- **Subsites**
- **Epidemiology**
- **Molecular Pathogenesis of Head and Neck Squamous Cell Carcinoma**
- **Human Papillomavirus–Associated Head and Neck Squamous Cell Carcinoma**

Current Treatment Paradigms
- **Unique Challenges of HNSCC**
- **Early Cancer**
- **Locoregionally Advanced Cancer**
- **Distantly Metastatic and Recurrent Disease**

Translational Approaches to Treatment of Head and Neck Cancer
- **Targeted Molecular Therapy**
- **Targeting Tumor Dormancy and Residual Disease in Head and Neck Squamous Cell Carcinoma**
- **Immunotherapy**
- **Robot-Assisted and Minimal-Access Surgery**

Conclusion
- **Case Presentation: Sunitinib + Image-Guided Radiation for Oligometastases**

Address Correspondence to:

Andrew Sikora
Mount Sinai School of Medicine
New York, NY
Email: andrew.sikora@mssm.edu

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20216

© 2010 Mount Sinai School of Medicine

ABSTRACT

Head and neck squamous cell carcinoma includes cancers of the mouth, throat, larynx, and lymph nodes of the neck. Although early disease is amenable to single-modality treatment with surgery or radiation, patients with advanced disease have a dramatically worse prognosis, despite potentially morbid/toxic treatment regimens involving surgery, radiation, chemotherapy, or all 3 modalities. The present review seeks to provide an overview of current understanding and treatment of head and neck squamous cell carcinoma for the nonspecialist clinician or basic/translational researcher, followed by an overview of major translational approaches to the treatment of head and neck squamous cell carcinoma. Translational research topics addressed include targeted molecular therapy, immunotherapy, minimally invasive robotic surgery, and ablation of dormant/residual tumor cells. Despite the many potentially promising avenues of head and neck squamous cell carcinoma research, only 2 new treatment approaches (antiepidermal growth factor receptor therapy and robotic surgery) have been approved by the US Food and Drug Administration in the past 30 years. Focused research programs involving integrated teams of clinicians, basic scientists, and translational clinician-researchers have the potential to accelerate discovery and change treatment paradigms for patients with head and neck cancer. ***Mt Sinai J Med 77:684–699, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: head and neck cancer, immunotherapy, minimally invasive surgery, squamous cell carcinoma, targeted therapy, translational research, transoral robotic surgery.

BIOLOGY AND EPIDEMIOLOGY OF HEAD AND NECK CANCER

Subsites

Embedded in a structurally complex region, head and neck squamous cell carcinoma (HNSCC) originates in the mucosa of 5 major anatomic subsites: the oral cavity, oropharynx, larynx, hypopharynx, and nasopharynx. Because nasopharyngeal carcinoma possesses a unique biology, clinical course, and treatment paradigm distinct from other subsites, it will be excluded from this review.

Epidemiology

Head and neck squamous cell carcinoma ranks as the sixth leading cause of cancer death worldwide, developing annually in approximately 550,000 individuals and responsible for approximately 304,000 deaths.[1,2] In the United States, 49,000 new cases are projected to occur in 2010, with an expected 11,000 deaths.[3] Despite the overall decreased incidence of HNSCC in the United States over the past 3 decades, researchers have observed a significant increase in the incidence of squamous cell malignancies of the oral tongue, base of tongue, and tonsil, particularly in young white patients aged 20–44 years–likely due to rising incidence of human papillomavirus (HPV)-associated HNSCC.[4]

Tobacco and alcohol are well-recognized risk factors for the development of HNSCC, which act synergistically.[5] A retrospective case-control study that examined risk factors for oral and pharyngeal cancer showed that, although heavy smoking was associated with an odds ratio (OR) of 20.7, and heavy drinking carried an OR of 4.9, the combination resulted in an OR of >48.[6] Chronic secondary tobacco exposure has also been linked to upper aerodigestive tract carcinomas, and in one meta-analysis was associated with an increased risk (OR: 1.55) of HNSCC.[7] Importantly, smoking cessation has been shown to inversely impact the risk of HNSCC development, with one study finding a 50% risk reduction if smoking cessation occurred by 50 years of age.[8] Another environmental risk factor outside the United States, particularly in Southeast Asia where there is a high incidence of oral-cavity cancers, is betel quid–a chewed preparation composed of areca nut, lime, and sometimes tobacco, all wrapped in betel leaf.[9,10] In contrast, the consumption of fruits and vegetables has been associated with a decreased risk of HNSCC development.[11]

Molecular Pathogenesis of Head and Neck Squamous Cell Carcinoma

Mucosal damage from chronic tobacco and alcohol exposure has been well characterized, both in terms of its clinicopathologic course and the underlying molecular derangements responsible for tumor development. Premalignant lesions, including leukoplakia and erythroplakia, progress to invasive carcinomas along a well-described pathologic sequence.[12] Molecular events that undergird this process include increasing cytogenic abnormalities, inactivation of tumor suppressor genes, and changes in intracellular signaling pathways that induce cellular immortalization.

A common genetic alteration in 70%–80% of dysplastic squamous cells and HNSCC tumors is the loss of chromosome 9p21, a region that contains CDKN2A encoding tumor suppressor genes p16 and p14 and is involved in G1 phase cell-cycle regulation.[12] Loss of 3p, a locus with tumor suppressor phenotype, is another common genetic event seen early in dysplasia.[12] Inactivation of tumor-suppressor genes through homozygous deletion, point mutations, and epigenetic alterations such hypermethylation fuel the neoplastic process. Inactivation of p16 is a frequent event witnessed in >80% of tumor specimens.[13] Similarly, loss of heterozygosity of 17p–the region encoding tumor suppressor p53–is extremely common.[13] Studies have found that half of all tumor specimens from patients with HNSCC contain p53 mutations. Notably, disruption of TP53–the genetic locus on 17p giving rise to p53–has been associated with reduced survival after surgical therapy for HNSCC.[14] Telomerase, an enzyme active in germ line cells but normally quiescent in somatic cells, has been shown to be overexpressed in 90% of HNSCC cells. Telomerase is responsible for maintaining genomic stability by protecting chromosomal ends, especially in rapidly dividing cells; its activity in malignant cells thereby enables evasion of apoptosis and contributes to cellular immortality.[15] Altered intracellular signaling also facilitates neoplastic development in HNSCC, including activation of oncogenic pathways downstream of the epidermal growth factor receptor (EGFR) and other molecular pathways. These will be explored in more detail in the appropriate sections below.

Human Papillomavirus–Associated Head and Neck Squamous Cell Carcinoma

Human papillomavirus is a recently appreciated cause of HNSCC.[16] A circular, double-stranded DNA virus, HPV encompasses many different subtypes, with HPV-16 and HPV-18 being the most common oncogenic variants in HNSCC. Using in-situ hybridization, HPV-16 DNA has been found in up to 72% of oropharyngeal cancer specimens where the association remains the highest.[17] Tumors related to HPV have also been linked to patterns of sexual activity, including a history of multiple sexual partners, early age of first intercourse, and frequent oral sexual practices.[17] On a molecular level, HPV gains access to the intracellular compartment of mucosal squamous cells and integrates into host DNA. The integrated virus subsequently expresses oncoproteins E6 and E7, which act synergistically to target the tumor suppressor genes p53 and pRb for ubiquitin-mediated intracellular degradation, resulting in genomic instability and oncogenic transformation as the normal cell-cycle regulatory points are inactivated.[16]

Human papillomavirus–related HNSCC clinically behaves differently than tobacco- and alcohol-induced HNSCC. Patients with HPV tend to present at a younger age, and without a history of excessive tobacco or alcohol use.[16] Overall, HPV+ HNSCC patients also have better outcomes, with tumors more responsive to both surgical and nonsurgical therapies and a lower risk of dying from disease. A recent retrospective analysis that included 323 patients with advanced oropharyngeal carcinoma who were prospectively enrolled in the Radiation Therapy Oncology Group (RTOG) 0129 trial examined the association between HPV status and survival. It showed that patients with HPV+ tumors, 63.8% of the group, had better 3-year survival (82.4% versus 57%); and, when adjusting for various factors–age, race, stage, tobacco and alcohol consumption, as well as treatment assignment–found a 58% reduction in the risk of death over a 4-year follow-up period.[18] Positive HPV status is an important prognosticator of response to therapy and survival, and development of oncologically effective treatment strategies that do not overtreat this population is an important focus of clinical and translational research.

> *Human papillomavirus is a recently appreciated cause of head and neck squamous cell carcinoma. Positive HPV status is an important prognosticator of response to therapy and survival, and development of oncologically effective treatment strategies that do not overtreat this population is an important focus of clinical and translational research.*

CURRENT TREATMENT PARADIGMS

Unique Challenges of HNSCC

Treatment requires a delicate balance between complete eradication of cancer and preservation of the underlying anatomical form and function. Thus, there are many challenges unique to the treatment of patients with HNSCC. Functionally, the head and neck region supports many fundamental physiologic processes including respiration, articulation, mastication, and deglutition.

> *Functionally, the head and neck region supports many fundamental physiologic processes including respiration, articulation, mastication, and deglutition.*

Tumors often significantly impact these processes through their involvement of vital structures such as cranial nerves, the carotid artery, skull base, and vertebral column. Eradicating the cancer while sparing surrounding structures–ie,organ preservation–assumes a paramount role in clinical decision-making. Quality of life is another critical factor in treatment planning, as structures of the head and neck form the basis for human social interactions.

Advances have been made on a variety of fronts to meet these challenges. In the case of surgical extirpation of locally advanced tumors, the development of microvascular free tissue transfer has enhanced reconstruction of postablative wounds by minimizing donor-site morbidity, restoring 3-dimensional soft-tissue characteristics of the defect, and carrying the potential for sensory innervation.[19] Other surgical advances include endoscopic and robotic techniques that currently supplement, and for some sites may ultimately replace, more morbid open surgical techniques.

Within the field of radiation oncology, intensity-modulated radiation therapy permits delivery of high-dose radiation to tumor sites while sparing vital radiation-sensitive surrounding structures such as the brain, spinal cord, and salivary glands. Similarly, medical oncology has emerged as a significant treatment modality for HNSCC. Once relegated to the domain of palliative management, chemotherapy has become an integral component of organ-preserving treatment strategies. Incorporation of chemotherapy into the treatment of locally advanced HNSCC has resulted in improved locoregional control and overall survival compared with treatment with radiation alone,[20] although at the cost of increasing toxicity. These treatment strategies are increasingly being incorporated into multimodality treatment schemes for both definitive treatment, and postoperative adjuvant therapy for high-risk patients.

Early Cancer

The American Joint Committee on Cancer outlines staging for the major subsites of the head and neck region and includes information regarding subsite, size of lesion (T), involvement of locoregional lymph nodes (N), and evidence of distant metastasis (M) Table 1.

Approximately one-third of patients will present with stage I and II disease. Early presentation of HNSCC is associated with a high cure rate, up to 90% for stage I disease and 70% for stage II disease. Single-modality therapy with either radiation or surgery usually is adequate and yields equal rates of locoregional control and survival. In general, the subsite involved by disease dictates the type of treatment most appropriate for the patient. Early oral-cavity lesions are treated surgically to avoid radiation toxicity to the mandible and salivary glands; laryngeal lesions are generally amenable to either radiation, transoral laser surgery, or (less frequently) open partial laryngectomy; and oropharyngeal cancer is treated with radiation therapy to avoid the substantial morbidity of open surgery in this area. However, transoral robotic surgery (TORS) is an emerging paradigm for treatment of selected cancers of the oropharynx and other head and neck subsites (see below). Early hypopharyngeal cancers are exceedingly rare, but are generally treated with radiation therapy.

Table 1. *AJCC Classification System for Head and Neck Squamous Cell Carcinoma.*

Stage	Tumor Size	Nodes	Metastasis
I	T1	N0	M0
II	T2	N0	M0
III	T3	N0	M0
	T1	N1	M0
	T2	N1	M0
	T3	N1	M0
IVA	T4a	N0	M0
	T4a	N1	M0
	T1	N2	M0
	T2	N2	M0
	T3	N2	M0
	T4a	N2	M0
IVB	Any T	N3	M0
	T4b	Any N	M0
IVC	Any T	Any N	M1

The American Joint Committee on Cancer (AJCC) outlines staging for the major subsites of the head and neck region and includes information regarding subsite, size of lesion (T), involvement of locoregional lymph nodes (N) and evidence of distant metastasis (M). The AJCC staging system is tightly correlated with overall survival and serves as an important role in prognostication. T1 and T2 lesions without cervical lymph node involvement represent early stage I and II disease respectively. T3 and T4 lesions denote progressively more advanced and locally infiltrative lesions. If no cervical disease is present, T3 and T4 lesions are characterized as Stage III and IV respectively. Nodal involvement automatically upstages a tumor despite its T size; N1 disease is categorized as Stage III disease, while N2-3 disease correlates with Stage IV disease. Distant metastatic disease, a rare presentation of HNSCC, automatically qualifies as Stage IV disease.

If surgical therapy is chosen for early disease management, an elective neck dissection can potentially provide essential histologic information to guide prognosis and further treatment decisions. In general, an elective neck dissection is performed if the risk of spread of disease from the primary site to the cervical nodal chains is >20% (which is generally the case for the oral cavity, oropharynx, hypopharynx, and supraglottic region of the larynx). High-risk histological features including perineural invasion, extracapsular spread, and multiple positive lymph nodes identify patients at risk for poor locoregional control and systemic disease after surgery. Phase III trials demonstrate that these patients benefit from adjuvant chemoradiation to improve locoregional control and disease-free survival (DFS).[21–23]

Locoregionally Advanced Cancer

Approximately two-thirds of patients will present with high T-stage or regional nodal metastasis, and thus are considered to have locoregionally advanced HNSCC, defined as stage III and IV. Advanced-stage disease significantly impacts survival, with 5-year survival rates ranging between 20% and 40% across the various subsites. Accordingly, multimodality treatment employing either concomitant chemoradiation or surgery followed by adjuvant chemoradiation is used. Numerous clinical studies demonstrate that both strategies are overall associated with equivalent rates of locoregional control and survival.[20,24–27]

For those patients who undergo nonsurgical treatment, concurrent cisplatin-based chemoradiation is the standard of care for advanced locoregional HNSCC. Unless a patient has significant comorbidities and poor performance status, chemotherapy should be included in the management of patients with advanced cancers. In the setting of unresectable disease, which includes invasion of the prevertebral musculature, carotid artery, and skull base, concomitant chemotherapy with radiation is the treatment of choice. Moreover, those sites associated with better functional recovery with nonsurgical management, including the oropharynx, hypopharynx, and larynx, are commonly managed with concurrent chemotherapy and definitive radiation.

Surgery followed by postoperative chemoradiation remains an important component of managing advanced, locally aggressive HNSCC. Advanced oral-cavity carcinomas are generally treated surgically to avoid cosmetic and functional deformities that would result from aggressive chemoradiation to this area. Surgical therapy is also generally preferred in patients with large, bulky laryngeal and hypopharyngeal tumors with cartilage invasion or severe baseline organ dysfunction, including paralyzed vocal cords, severe dysphagia requiring pretreatment gastrostomy tube placement, or intractable aspiration. In these cases, organ-preservation strategies would render the patient anatomically intact, but with a dysfunctional larynx.[24,25,27] Lastly, surgical extirpation is a consideration in patients who have failed conservation strategies and still have resectable disease, to obtain ultimate locoregional control.

Distantly Metastatic and Recurrent Disease

Despite aggressive multimodality therapy employed in the management of locally advanced disease, distant metastases develop in approximately 20%–30% of patients.[27] With the exception of focally distant, resectable lesions such as pulmonary metastases, distant disease is classically considered incurable and treatment regimes focus on palliation. Chemotherapeutic agents are the mainstay of treatment and include standard cytotoxic agents alone or in combination, as well as novel targeted biological agents. Combination therapy with standard cytotoxic agents, including platinum-based compounds, taxanes, 5-flurouracil, methotrexate, and ifosfamide, generally results in response rates of 20%–40% with duration of benefit averaging 2–4 months, without significant improvement in overall survival when compared with single-agent therapy.[28,29] The development of novel therapeutic agents is challenging the current treatment paradigm for distant metastatic disease, a topic that will be highlighted in the following sections.

Recurrent disease, however, is often amenable to treatment. Approximately half of those with locally advanced, HPV− HNSCC will relapse locoregionally within 2 years despite aggressive regimens that include concurrent chemoradiation and/or surgery.[18,22] In the setting of locoregional recurrence, complete surgical resection is the treatment of choice, but it is often not possible or incomplete.[30] In a randomized trial reported by the European Organisation for Research and Treatment of Cancer, approximately 80% of patients who underwent salvage surgery for recurrent disease following curative radiation suffered locoregional relapse within 24 months.[30] In contrast, reirradiation, which has been extensively explored in the management of recurrent HNSCC, offers improved locoregional control and DFS. The 3-year locoregional control following salvage surgery and adjuvant chemo-reirradiation is approximately 60%.[31] For unresectable patients, the 3-year locoregional control following reirradiation and concurrent chemotherapy is approximately 35%.[31] Due to the competing risks of locoregional relapse, distant relapse, and severe complications following reirradiation, only 20% to 30% remain alive and free of disease at 3 years following treatment.[30,31] Therefore, this is an appropriate population with which to conduct translational trials investigating novel radiosensitizing agents with the goal of increasing efficacy and reducing toxicity.

TRANSLATIONAL APPROACHES TO TREATMENT OF HEAD AND NECK CANCER

Targeted Molecular Therapy

The identification of specific molecules and pathways that drive the growth and survival of HNSCC provides opportunities for development of biologically based therapies that preferentially target cancer cells. Although thus far only agents targeting the epidermal growth factor receptor (EGFR) pathway have entered routine clinical use, a number of other promising approaches have been identified and are the subject of ongoing research.

Agents Targeting the Epidermal Growth Factor Receptor

Epidermal growth factor receptor, a member of the ErbB growth factor receptor tyrosine kinase family, is upregulated in >90% of HNSCC cases and is associated with poor prognosis, according to many retrospective analyses.[32,33] Binding to EGFR by any of its molecular ligands, most notably tumor growth factor-α (TGF-α) and EGF, results in receptor homodimerization or heterodimerization with other members of the ErbB family. Once bound, EGFR activates its associated tyrosine kinase, which leads to downstream intracellular signal transduction cascades involving signal transducer and activator of transcription (STAT), mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3K), and phospholipase C (PLC) pathways, controlling cellular proliferation, apoptosis, angiogenesis, and the propensity for invasion.[34] Upregulation of EGFR can result from either gene amplification or transcriptional activation, and is nearly ubiquitous in HSNCC.[35,36] Thus, inhibition of EGFR signaling, either by interrupting the binding of its activating ligands to the extracellular domain (the mechanism of the chimeric human/mouse antibody cetuximab and the fully humanized antibody panitumumab) or interfering with signal transduction mediated by the intracellular (kinase) domain (the mechanism of the kinase inhibitors gefitinib and erlotinib), is a rational therapeutic approach.

Cetuximab was shown to have single-agent efficacy in patients with chemotherapy-resistant HNSCC in a multicenter phase II trial[37] and significantly enhanced locoregional control in combination with radiation.[38] These studies led to the first regulatory approval, in 2006, of a new agent for HNSCC in decades. Gefitinib and erlotinib have also been shown to have efficacy alone and in combination with radiation and chemotherapy.[39,40] Overall, toxicity profiles of anti-EGFR therapies have compared favorably with chemotherapeutic regimens. Thus, anti-EGFR therapies have entered routine clinical practice, particularly in patients who have failed prior chemotherapy or cannot medically tolerate conventional cytotoxic chemotherapy.

Although EGFR blockade has been shown to be clinically effective, overall results are modest and most patients do not show significant benefit, as demonstrated by major response rates on the order of 10%–20%. Thus, even in the absence of prior therapy with anti-EGFR agents, the majority of patients display intrinsic treatment resistance. Furthermore, patients who respond initially often develop acquired resistance to EGFR inhibitors, thus limiting their long-term efficacy. Interestingly, resistance does not seem to correlate with EGFR expression levels,[41] and in the case of cetuximab does not correlate with tyrosine kinase (signaling) domain mutations. For the tyrosine kinase inhibitors (TKIs) gefitinib and erlotinib, there is some evidence that kinase domain mutations, although not commonly identified, correlate with efficacy, and this is an active topic of research.[42] Other mechanisms of resistance described include overproduction of other growth factors, such as vascular endothelial growth factor (VEGF),[43] and activation of overlapping or parallel signal transduction pathways, such as the c-MET pathway.[43] Because innate and acquired resistance to anti-EGFR agents is a major barrier to clinical efficacy, development of strategies for reversing resistance, or selecting those patients most likely to respond, is a major research focus.

> *Because innate and acquired resistance to anti–epidermal growth factor receptor agents is a major barrier to clinical efficacy, development of strategies for reversing resistance, or selecting those patients most likely to respond, is a major research focus.*

Antivascular and Other Ligand/Receptor/Kinase-Targeted Approaches

The relative success of agents targeting EGFR has led to interest in targeting other HNSCC growth/survival pathways. Angiogenesis is critical to tumor development, as unregulated, rapidly dividing cells require a constant nutritional support supplied largely by the vasculature. Angiogenesis is promoted by a number of tumor-derived mediators, including VEGF, basic fibroblast growth factor (bFGF), prostaglandin E_2 (PGE_2), transforming growth factor-β (TGFβ), interleukin-8 (IL-8), and angiopoietin-1.[44] Studies examining aerodigestive-tract carcinomas have shown that HNSCC cells release PGE_2 and TGFβ, which in turn stimulate endothelial-cell motility.[45] Because antiangiogenic therapy targeting VEGF and its receptors has been used in the treatment of other cancers, this approach has also been tried, alone and in combination with EGFR inhibition, in the treatment of HNSCC. A family of related molecules (VEGF-a, -b, and -c), VEGF has been shown to be overexpressed by HNSCC, to promote both lymphatic and distant metastasis, and to be correlated with poor prognosis.[46] VEGF overexpression has also been described as a mechanism of therapeutic resistance to EGFR inhibitors.[43] Thus, the anti-VEGF antibody bevicizumab has been tested in combination with erlotinib in a phase I/II trial that included both previously untreated HNSCC patients and patients who failed prior therapy.[47] There were some major responses, including several patients who had complete regression of tumor, but results of combination therapy were only better for previously untreated patients, and there were significant bleeding complications.

The HER2 and HER3 receptors are capable of dimerizing with the EGFR and activating the Akt/PI3K/mammalian target of rapamycin (mTOR) growth/survival pathway, and HER2 has been validated as a clinical target in breast cancer and other solid tumors.[48] Specifically in HNSCC, EGFR upregulation induced by chronic exposure to cetuximab has been shown to promote dimerization with HER2 and HER3 and activate Akt/PI3K.[49] This was blocked by concomitant HER2 and EGFR inhibition, thus providing a rationale for combination molecular therapy with clinically available EGFR and HER2 inhibitors, as well as development of molecules to inhibit HER3. Although less well studied in HNSCC, specific targets of the mTOR pathway (inhibited by rapamycin, and second-generation mTOR inhibitors currently in clinical and preclinical trials) represent additional areas for further investigation.[50] Additional oncogenic growth/survival pathways, although less well described in HNSCC, also provide potential targets for treatment, as they are inhibited by molecules currently under intensive clinical or preclinical development. These include the insulin-like growth factor 1 receptor (IGF-1R) pathway[51] and the Aurora kinase regulators of mitosis.[52]

So-called "multitarget" TKIs, which inhibit multiple growth factor receptors with varying degrees of efficacy, have also been tested in HNSCC, particularly those with activity against EGFR and/or VEGF. Sorafenib is a highly promiscuous TKI with activity against VEGFR, platelet-derived growth factors (PDGFR), c-KIT, flt-3, and the threonine-serine kinase RAF. Sorafenib has been tested in several trials for recurrent/metastatic HNSCC; it appears to have activity against treatment-naïve HNSCC but little activity against heavily pretreated disease.[53,54]

The standard approach to head and neck cancer with distant metastasis is palliative chemotherapy with best supportive care.[55] Sunitinib is another multitarget TKI, with an activity profile similar to sorafenib. Our group is conducting a phase I/II trial investigating concurrent sunitinib and image-guided radiation therapy followed by maintenance sunitinib for patients with oligometastases.[56] The rationale for administering curative-intent radiation for oligometastases is that a proportion of these patients has been shown to be rendered free of detectable disease with an acceptable toxicity profile.[57] In these studies, approximately 20% of patients remain free of recurrence several years after radiation.[57,58] As expected, most patients develop additional distant metastases within months of treatment.[57,58] This highlights the need for effective systemic agents for the majority of patients. To date, we have treated 38 patients with this approach, including 6 patients with metastatic HNSCC. At a median follow-up of 17 months, the 18-month overall survival, local control, distant control, and progression-free survival were 50%, 74%, 46%, and 28%, respectively. To date, there was no toxicity ≥ grade 3 related to radiation.[56] Oligometastases should now be considered a potentially curable state in selected head and neck cancer patients.[59]

Oligometastases should now be considered a potentially curable state in selected head and neck cancer patients.

Targeting Cancer-Associated Inflammation

The relationship between cancer and inflammation has been known since antiquity, but only recently have specific molecular determinants been identified as potential therapeutic targets. Chronic inflammation is a well-known contributor to many cancers (liver, colon, esophageal) and has been proposed as a factor in HNSCC carcinogenesis, particularly in the case of oral cancer.[60] Conversely, established tumors create an inflammatory milieu through induction of necrosis, generation of reactive oxygen and nitrogen intermediates, and expression of proinflammatory cytokines and other soluble mediators. Thus, inactivating or inhibiting proinflammatory molecules can potentially suppress carcinogenesis and tumor progression, and promote regression of established tumors.

One well-known inflammatory mediator that has been targeted in HNSCC is the cyclooxygenase-2 (COX-2) enzyme. Products of COX-2 are pleiotropic modulators of inflammation and diverse cellular functions, including growth, apoptosis, and cell-cell signaling. In cancer cells, COX-2 upregulation has been implicated in HNSCC carcinogenesis, progression, and treatment resistance.[61] Inhibition of COX-2, usually in combination with radiotherapy and chemotherapy or other targeted agents, has been intensively investigated as a chemopreventive agent[61] as well as treatment for recurrent and metastatic disease. For instance, our group recently completed a phase I trial of concurrent erlotinib, celecoxib, and reirradiation for patients with recurrent head and neck cancer.[62] This represents a "first in man" clinical trial of combining EGFR and COX-2 inhibition with radiotherapy, which was supported by strong preclinical rationale and laboratory data.[63,64] This regimen was feasible and well tolerated even in heavily pretreated patients with recurrent HNSCC that progressed despite previous radiation.[62] At a median follow-up of 15 months, 4 of 14 treated patients remain alive and free of progression.[62] Further investigation of novel agents targeting cell signaling, angiogenesis, and immune cells in combination with reirradiation are ongoing.

Inducible nitric oxide synthase (iNOS, NOS2) is another inflammatory molecule overexpressed in HNSCC and potentially amenable to pharmacologic therapy. The enzyme iNOS and its product, nitric oxide (NO), are implicated in many aspects of cancer pathobiology, including angiogenesis, metastasis, and immune evasion.[65,66]

> *The enzyme inducible nitric oxide synthase and its product, nitric oxide, are implicated in many aspects of cancer pathobiology, including angiogenesis, metastasis, and immune evasion.*

Intratumoral iNOS expression predicts poor clinical prognosis in a number of cancers, including melanoma and HNSCCA. Expression of iNOS in HNSCCA has been shown to be clinically correlated with angiogenesis,[67] extracapsular spread,[68] and an increased rate of lymphatic metastasis. Thus, there is substantial clinical evidence supporting a role for iNOS and NO in the pathobiology of HNSCCA.

Many small-molecule iNOS inhibitors exist,[69] and some have been previously tested in human clinical trials of benign disease. Inhibition of iNOS, alone or in combination with cytotoxic chemotherapy, has been tested as a therapeutic approach in several preclinical tumor models, most extensively in melanoma. Using a preclinical model of human melanoma, members of our group have demonstrated synergistic tumor control without additive toxicity when the selective iNOS inhibitor L-NIL (L-N6- (1-iminoethyl)-L-lysine hydrochloride) was combined with cisplatin chemotherapy.[70] We have now generated preliminary evidence that targeted iNOS inhibition has activity against human HNSCC xenografts as a single agent, and we are currently testing its efficacy in combination with cisplatin. Another group has shown that a nonspecific nitric oxide synthase inhibitor (L-NAME) could diminish proliferation of a HNSCCA cell line in vitro.[71] These observations have exciting implications for HNSCC, where cisplatin remains a first-line chemotherapeutic agent, and suggest that targeted iNOS inhibition is a promising approach to treatment of HNSCC but requires further confirmation in preclinical models.

Targeting Tumor Dormancy and Residual Disease in Head and Neck Squamous Cell Carcinoma

Following surgical resection of the primary tumor, even if the margins are free of cancer, >50% of patients develop locoregional or distant organ recurrences.[72] To prevent these, it is crucial to understand the biology of residual disease so that it can be targeted before progression to overt lesions. Tumor dormancy describes the ability of HNSCC cells to enter a state of low proliferation, and thus persist undetected. Our definition of tumor-cell dormancy is quiescence,[73] an uncommitted and reversible growth arrest that can be brought about by different signals[74] that can prevent cancer cells from entering differentiation or irreversible senescence,[74] 2 cellular endpoints that tumor cells can evade.[73] The fact that in patients with recurrent disease tumor dormancy is "reversible" further supports the definition of quiescence. The mechanisms that determine whether tumor cells, after disseminating to target organs, will continue to proliferate, die, or enter a protracted state of dormancy are largely unknown.[73]

> *The mechanisms that determine whether tumor cells, after disseminating to target organs, will continue to proliferate, die, or enter a protracted state of dormancy are largely unknown.*

Thus, it is of utmost importance to determine the molecular-genetic mechanisms driving dormant tumor cell (DTC) biology,[73] including the poorly understood mechanisms driving growth arrest and prolonged dormancy of DTCs.[72,75,76] Thus, it has been proposed that DTCs that survive treatment may remain in a dormant state. In fact, it is known that growth-arrested cells can evade certain chemotherapeutic regimens. This is not only because the cells are not cycling, but our published evidence suggests that active mechanisms dependent on the activation of a stress-adaptive unfolded protein response, or UPR, plays an active role in protecting cancer cells from chemotherapy-induced stress.[73] It is clear that targeting and eradicating these DTCs would represent a significant clinical breakthrough.

Our group has shown that activation of the kinase p38, while inhibiting extracellular-signal-regulated kinase (ERK) signaling in the human HNSCC line HEp3, activates the UPR, which allows dormant HEp3 cells to survive nutritional stress insults as well as damage induced by common chemotherapy drugs such as doxorubicin, etoposide, and cisplatin. Activation of p38 also induces the transcriptional upregulation of p53 (R213Q mutant) and basic helix-loop-helix family (BHLHB3) transcription factors, which are responsible for the induction of quiescence. The $p53^{mutR213Q}$ protein appears to retain the capacity to induce growth arrest but not apoptosis.[77] When compared with other mutations in p53 in oral squamous carcinoma patients, those bearing the R213Q mutation have longer recurrence-free periods and overall survival.[78] We propose that the presence of the R213Q p53 mutant protein (or other mutations with a similar function) may predispose DTCs to become dormant and avoid apoptosis imposed by the target-organ microenvironment and/or therapy. This mutation, and perhaps others that promote quiescence, may allow DTCs to most likely remain undetected, explaining the more favorable prognosis.

The translational impact of addressing the problem of minimal residual disease and DTCs in HNSCC is high. It is commonly accepted that metastases can originate from DTCs, yet the mechanisms responsible for this transition are mostly unknown. In the case of HNSCC, local recurrences might also originate from locally DTCs that escape detection. Because eventual success of anticancer therapy depends strictly on curing or preventing metastases, progress in this field is urgently needed to improve the control of metastatic precursors by maintaining dormancy as a chronic asymptomatic condition, or eradicating dormant tumor cells.

Immunotherapy

Similar to other malignancies, head and neck carcinomas rely on immune evasion to escape host immune defense. This includes both direct inhibition and manipulation of immune surveillance. Studies of patients with HNSCC have shown that these patients, compared with normal controls, have reduced peripheral blood concentrations of CD3+, CD4+, and CD8+ cells, and that these differences can persist even after treatment.[79] Moreover, in-vitro studies have shown that HNSCC tumor cells are capable of downregulating presentation of tumor antigen via major histocompatibility complex class I molecules, thereby evading immune recognition and destruction.[80,81] Because the ability of HNSCC to persist in the host represents a failure of antitumor immunity, approaches have been developed to boost HNSCC-directed immune responses, either by stimulating T-cell–dependent immunity or blocking immunoregulatory pathways that suppress antitumor responses. The ultimate goal of most of these approaches is to generate a strong CD8+ cytotoxic lymphocyte (CTL) response capable of eradicating the tumor and providing long-term immunologic memory.

Tumor Vaccination

Although there are many different strategies for antitumor vaccination, all share a number of common characteristics[70]: a target antigen, or antigens; an adjuvant to enhance the immune response; and a delivery system. Vaccine strategies usually focus on tumor-associated antigens (TAA), although other strategies, such as immunization against proteins overexpressed in angiogenesis,[82] have been proposed. Tumor-associated antigens in HNSCC can be unique to the tumor (such as E6, E7, and other HPV proteins); can be mutated in the tumor (such as p53); can be overexpressed by the tumor (p53, EGFR); or can be lineage-restricted (melanoma-associated MAGE, cancer/testis NY-ESO-1).[83] Antigens delivered in vaccines can take the form of proteins, immunogenic peptides, cell lysates, or whole killed/irradiated cells. Adjuvants can be simple, such as incomplete Freund's adjuvant or alum (which both serve as antigen depots at the site of vaccination); or more complex, such as coadministration of toll-like receptor (TLR) ligands (such as CpG DNA or imiquimod), cytokines (such as IL-12 or granulocyte-macrophage colony stimulating factor [GM-CSF]), or co-injection of dendritic cells (which are professional antigen-presenting cells). Delivery systems include subcutaneous injection, DNA vaccination (in which the antigen is expressed within host cells), genetically modified microorganisms, and even more esoteric strategies such as the creation of virus-like particles. Many combinations of the above have been tested in clinical trials of HNSCC,[83,84] but vaccine trials have tended to focus on generating a response to p53 (due to the high rate or mutation/overexpression in HNSCC) or HPV viral proteins, particularly E6 and E7 (which are foreign antigens, as well as essential for maintaining the malignant phenotype). As has been the case with vaccine clinical trials in many other solid tumors, induction of vaccine-specific cytotoxic T lymphocytes (CTL) is frequently seen, but clinical responses are rare. As clinical trials shift toward immunizing in the adjuvant setting, rather than at advanced stages of disease, it will be possible to assess the impact of vaccination on less-immunocompromised patients, and evaluate the impact of vaccination on other endpoints, such as recurrence and incidence of second-primary tumors.

As clinical trials shift toward immunizing in the adjuvant setting, rather than at advanced stages of disease, it will be possible to assess the impact of vaccination on less-immunocompromised patients.

Adoptive Immunotherapy

Vaccination seeks to expand the patient's own circulating T-cell pool in vivo; adoptive immunotherapy involves expanding antigen-specific CTL (or other immune cell types) outside the body and reinfusing them into the patient. High numbers of antigen-specific CTL can be obtained by multiple rounds of ex-vivo stimulation of tumor infiltrating lymphocytes (TIL) with antigen, or introducing antigen-specific T-cell receptors into bulk lymphocytes through genetic engineering. Although this approach has been utilized extensively in other solid tumors (such as melanoma), and despite promising preclinical evidence of its efficacy in HNSCC, only 1 clinical trial of 15 HNSCC patients has been reported.[85] A variant of adoptive immunotherapy, in which peripheral blood leukocytes are stimulated ex vivo with IL-2 to produce so-called LAK cells (lymphokine-activated killer cells–a heterogeneous population of activated T and natural killer [NK] cells), has been tested clinically in HNSCC.[86] Despite sporadic reports of clinical response, this approach has mostly fallen out of favor. One major barrier to expansion of adequate numbers of antigen-specific cells from TIL in HNSCC is thought to be T-cell inactivation by the tolerogenic tumor microenvironment. Thus, the future of adoptive immunotherapy for HNSCC likely will depend on the creative application of strategies for reversal of tumor-mediated immunosuppression, or creation of genetically engineered tumor-specific T cells.

Reversal of Tumor-Mediated Immunosuppression

The fact that many HNSCCs are briskly infiltrated by lymphocytes, despite the difficulty of isolating functional T cells from tumor, highlights the important role that tumor-mediated immunosuppression (TMI) plays in HNSCC immune escape. Like most cancers, HNSCC actively suppress the immune system through release of soluble mediators and induction of suppressive/regulatory immunocytes.[87–89] Tumors can directly produce soluble mediators of immunosuppression, such as TGF-β, IL-10, and VEGF,[87,88] or can trigger their release by provoking an inflammatory response. These mediators can act directly on T cells or other immunocytes to induce tolerance and immunosuppression, or indirectly by induction of regulatory/suppressive immunocytes. Such cellular mediators of TMI include regulatory T cells (Treg) and myeloid-derived suppressor cells (MDSC). Treg, which express CD4, CD25 (IL-2 receptor), and the transcription factor forkhead box P3 (FOXP3), can suppress effector T-cell responses in both an antigen-specific and "bystander" fashion and can be induced by TGF-β.[87,90] Treg are often found at high levels in the tumor, tumor-draining lymph nodes, and blood of cancer patients.[90]

Myeloid-derived suppressor cells are immature myeloid cells that originate in the bone marrow and whose formation and release into the periphery is stimulated by the release of soluble mediators such as VEGF, GM-CSF, IL-1β, and a number of other soluble mediators released by tumors.[89,91] They are characterized by their potent ability to suppress T and NK cell function via increased expression of arginase I and iNOS, and increased production of reactive oxygen species.[92] Additionally, MDSC can promote the development of Treg.[93] Recently, there has been growing interest in targeting these MDSC in head and neck cancers with pharmacological agents including phosphodiesterase 5 inhibitors (sildenafil, tadalafil), COX-2 inhibitors (celecoxib), or c-kit inhibitors[94] (sunitinib, dasatinib).[56,62,95] Inhibition or ablation of immunosuppressive MDSC and/or Treg to unmask the endogenous immune response is a promising approach to reversal of TMI in HNSCC and other solid tumors that not yet been formally tested in clinical trials.

Another approach to boosting endogenous antitumor immune responses is blocking antibody (ipilumumab) to the molecule CTLA-4, a negative regulator of T-cell activation.[96] Ipilumumab has been shown to cause immune-mediated regression of melanoma and other solid tumors, and was recently shown to improve survival in melanoma.[97] The immune response to HNSCC, characterized as it is by the induction of plentiful, but dysfunctional, tumor-infiltrating lymphocytes, suggests that blockade of CTLA-4 and other immunoregulatory molecules is a logical approach to reversing TMI.

Antibody-Mediated Immune Stimulation

Whereas cetuximab and other anti-EGFR antibodies were developed as inhibitors of EGFR stimulation and downstream signal transduction, as antibodies they can potentially induce immunological effects as well.[98]

> *Whereas cetuximab and other anti-EGFR antibodies were developed as inhibitors of EGFR stimulation and downstream signal transduction, as antibodies they can potentially induce immunological effects as well.*

In fact, this appears to be the case, as in vitro studies demonstrate that cetuximab only induces lysis of HNSCC cell lines in the presence of donor leukocytes, suggesting that antibody-dependent cellular cytotoxicity via interactions with Fc receptors on innate immune cells may play a role in its antitumor effects.[99] Recent evidence that these mechanisms may be clinically relevant comes from studies in colorectal carcinoma[100] that demonstrate that patient responses to cetuximab correlate with polymorphisms in the Fcγ receptors, which mediate antibody binding to phagocytes and NK cells involved in antibody-dependent cellular cytotoxicity. Thus, it is likely that the clinical efficacy of anti-EGFR antibodies derives, at least in part, from their ability to stimulate the immune system. This is potentially an important observation that may have relevance to overcoming treatment resistance and integration of anti-EGFR antibodies into other immunotherapeutic strategies for HNSCC.

Robot-Assisted and Minimal-Access Surgery

Head and neck oncologic surgery has evolved as technological advances have led to the development of new surgical procedures that are organ-preserving and oncologically sound and significantly reduce patient morbidity and mortality. One of the latest is the application of TORS to malignancies of the upper aerodigestive tract, including HNSCC. Although robotic techniques have long been applied to cardiothoracic and abdominopelvic surgeries,[101–103] robotic-assisted surgery in head and neck cancer is a relatively recent development that has the potential to significantly impact treatment of HNSCC.[104,105] Although TORS may ultimately have broad applications to HNSCC, it is particularly well adapted to oropharyngeal cancer.[104] Open procedures for surgical exposure of the oropharynx often involve a lip-split and division of the mandible, even for the earliest malignant lesions. The resultant morbidity was unacceptable, and led to the paradigm shift toward nonsurgical organ-preservation strategies once proven to have survival and locoregional control rates equivalent to surgery.[21–23] Despite these trends, treatment with definitive chemoradiation can result in significant acute and late toxicity, and in one study resulted in 4% mortality with roughly 40% of patients exhibiting permanent gastrostomy-tube dependency, tracheotomy dependency, or need for salvage surgical procedures for failed locoregional control.[106]

It is in this context that the minimally invasive surgical procedures such as TORS developed. Seminal preclinical[105,107–109] and human phase II studies[110–112] using the da Vinci Surgical System have laid the groundwork by establishing the feasibility, safety, and efficacy of TORS, and similar work at our institution has expanded on the applicability of robotic-assisted surgery in managing malignancies of the head and neck.[113–116] Using the da Vinci robot, TORS involves a remote surgical console operated by the surgeon, and a surgical robotic cart positioned near the patient with 3 arms inserted transorally to access the surgical site. The 3 arms include a 3-dimensional endoscopic camera, as well as 2 laterally placed arms that hold microsurgical instruments and can be manipulated 360° to tightly mirror the movements of the operating surgeon.

A study at our institution that included 20 patients with early head and neck cancer at various upper–aerodigestive tract subsites showed that TORS provided excellent surgical exposure for resection, was able to achieve complete negative margins, and displayed no intraoperative or significant postoperative complications such as death, pneumonia, aspiration, hemorrhage, or airway obstruction.[115] Intraoperative blood loss was low, averaging 80 mL per procedure. Primary reconstruction of ablative defects was accomplished with TORS during the primary procedure. In 2 patients, access was deemed inadequate and the robotic procedure was terminated. However, all operated patients were extubated after the procedure and subsequently did not require tracheotomy. Most patients were immediately started on an oral diet after endoscopic evaluation of swallow examination and were discharged home on average in 1.7 days. Postsurgical adjuvant treatment was offered based on high-risk histologic features as detailed above. Although long-term results are pending, our cohort of patients remains free of recurrence and is not gastrostomy-tube or tracheotomy-tube dependent.

Because minimal-access surgery depends critically on obtaining negative margins, our group is actively working to integrate an advanced miniaturized fiberoptic microscope developed at Rice University into the robotic platform to allow real-time histological margin assessment. Preliminary ex vivo

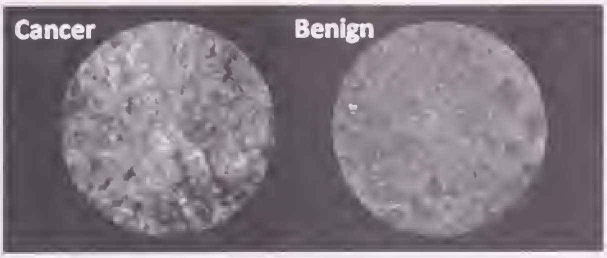


Fig 1. Representative oropharyngeal squamous cell carcinoma specimens imaged with proflavin-enhanced HRME immediately after resection with the da Vinci Surgical System. Note the irregular, crowded, and disorganized pattern of staining of the mucosal tumor (left) as compared with the organized nuclear pattern seen in the surrounding margin of normal mucosa (right). **Abbreviations:** HRME, high-resolution magnification endoscopy.

results have demonstrated that the majority of HNSCC specimens can be successfully imaged, and show good correlation with gold-standard histopathology. Future studies will test the ability of this technique to discriminate between benign and cancerous mucosa in vivo (Figure 1).

Although TORS has yet to be compared with definitive chemoradiation in a phase III clinical trial, the results from our experience and those from other institutions are extremely encouraging.

> *Although transoral robotic surgery has yet to be compared with definitive chemoradiation in a phase III clinical trial, the results so far are extremely encouraging.*

Supplanting the older, morbid, open surgical procedures, TORS is safe and yields excellent functional outcomes that may potentially exceed those of current nonsurgical organ-preserving strategies. Combining TORS with adjuvant therapy and assessing long-term outcomes in phase III studies against concurrent chemoradiation will provide more rigorous evaluation of the role of TORS in HNSCC.

CONCLUSION

For patients with head and neck cancer, the past 30 years have led to significant gains in esthetic/functional outcomes and quality of life, as reconstructive techniques improved and multimodality therapy became integrated into head and neck surgical oncology. However, improvements in disease control and survival have been modest at best. Thus there is a compelling unmet need for new treatment approaches for HNSCC. The recent integration of targeted molecular therapy (anti-EGFR agents) into routine therapy of head and neck cancer patients confirms the potential for translational research to alter patient care. A true translational approach, which thrives at the intersection of careful clinical observation and creative experimentation, will fuel continued advances in the field of head and neck oncology.

Case Presentation: Sunitinib + Image-Guided Radiation for Oligometastases

A 63-year-old man originally had locally advanced larynx cancer treated with chemoradiation at an outside hospital. He was found to have recurrent disease 5 months after completing radiation, and subsequently underwent salvage laryngectomy of an rT4 tumor with negative margins. However, 6 months after surgery, the patient was found to have 2 lung metastases, which were biopsy proven. The patient underwent salvage chemotherapy consisting of 5-fluorouracil, taxotere, and cetuximab for 3 cycles. A computed tomography scan demonstrated local progression and the development of a third lung metastasis. The patient was enrolled in the phase II trial of sunitinib (37.5 mg daily, days 1–28) and image-guided radiation (50 Gy in 10 fractions starting d8). He received radiation to all 3 lung lesions. The patient tolerated treatment well and continues on maintenance sunitinib (50 mg daily; 4 weeks on, 2 weeks off). Positron-emission tomography/computed tomography performed 3 months after radiation demonstrated no active disease (Figure 2).

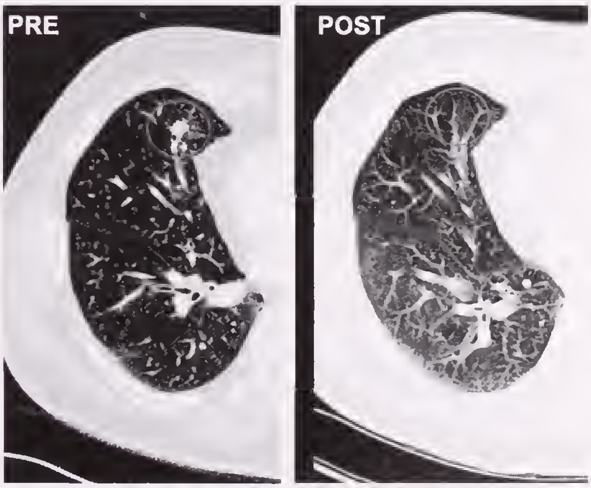


Fig 2. Chest CT scan demonstrating lung metastasis prior to treatment (pre) and resolution of the metastasis after sunitinib + image-guided radiation (post). **Abbreviations:** CT, computed tomography. This figure is available in color online at Wiley Online Library (wileyonlinelibrary.com).

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Parkin DM, Bray F, Ferlay J, et al. Global cancer statistics, 2002. *CA Cancer J Clin* 2005; 55: 74–108.
2. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2010; doi:10.1002/ijc.25516.
3. Jemal A, Siegel R, Xu J, et al. Cancer statistics, 2010. *CA Cancer J Clin* 2010; 60: 277–300.
4. Shiboski CH, Schmidt BL, Jordan RC. Tongue and tonsil carcinoma: increasing trends in the U.S. population ages 20–44 years. *Cancer* 2005; 103: 1843–1849.
5. Vineis P, Alavanja M, Buffler P, et al. Tobacco and cancer: recent epidemiological evidence. *J Natl Cancer Inst* 2004; 96: 99–106.
6. Rodriguez T, Altieri A, Chatenoud L, et al. Risk factors for oral and pharyngeal cancer in young adults. *Oral Oncol* 2004; 40: 207–213.
7. Lee YC, Boffetta P, Sturgis EM, et al. Involuntary smoking and head and neck cancer risk: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Cancer Epidemiol Biomarkers Prev* 2008; 17: 1974–1981.
8. Bosetti C, Gallus S, Peto R, et al. Tobacco smoking, smoking cessation, and cumulative risk of upper aerodigestive tract cancers. *Am J Epidemiol* 2008; 167: 468–473.
9. Warnakulasuriya S. Areca nut use following migration and its consequences. *Addict Biol* 2002; 7: 127–132.
10. Trivedy CR, Craig G, Warnakulasuriya S. The oral health consequences of chewing areca nut. *Addict Biol* 2002; 7: 115–125.
11. Freedman ND, Park Y, Subar AF, et al. Fruit and vegetable intake and head and neck cancer risk in a large United States prospective cohort study. *Int J Cancer* 2008; 122: 2330–2336.
12. Perez-Ordoñez B, Beauchemin M, Jordan RC. Molecular biology of squamous cell carcinoma of the head and neck. *J Clin Pathol* 2006; 59: 445–453.
13. Reed AL, Califano J, Cairns P, et al. High frequency of p16 (CDKN2/MTS-1/INK4A) inactivation in head and neck squamous cell carcinoma. *Cancer Res* 1996; 56: 3630–3633.
14. Poeta ML, Manola J, Goldwasser MA, et al. TP53 mutations and survival in squamous-cell carcinoma of the head and neck. *N Engl J Med* 2007; 357: 2552–2561.
15. McCaul JA, Gordon KE, Clark LJ, et al. Telomerase inhibition and the future management of head-and-neck cancer. *Lancet Oncol* 2002; 3: 280–288.
16. Chung CH, Gillison ML. Human papillomavirus in head and neck cancer: its role in pathogenesis and clinical implications. *Clin Cancer Res* 2009; 15: 6758–6762.
17. D'Souza G, Kreimer AR, Viscidi R, et al. Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl J Med* 2007; 356: 1944–1956.
18. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010; 363: 24–35.
19. Genden EM. Reconstruction of the mandible and the maxilla: the evolution of surgical technique. *Arch Facial Plast Surg* 2010; 12: 87–90.
20. Pignon JP, le Maître A, Maillard E, et al. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. *Radiother Oncol* 2009; 92: 4–14.
21. Bernier J, Vermorken JB, Koch WM. Adjuvant therapy in patients with resected poor-risk head and neck cancer. *J Clin Oncol* 2006; 24: 2629–2635.
22. Cooper JS, Pajak TF, Forastiere AA, et al. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. *N Engl J Med* 2004; 350: 1937–1944.
23. Bernier J, Domenge C, Ozsahin M, et al. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. *N Engl J Med* 2004; 350: 1945–1952.
24. Weber RS, Berkey BA, Forastiere A, et al. Outcome of salvage total laryngectomy following organ preservation therapy: the Radiation Therapy Oncology Group trial 91–11. *Arch Otolaryngol Head Neck Surg* 2003; 129: 44–49.
25. The Department of Veterans Affairs Laryngeal Cancer Study Group. Induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. *N Engl J Med* 1991; 324: 1685–1690.
26. Lefebvre JL, Chevalier D, Luboinski B, et al.; EORTC Head and Neck Cancer Cooperative Group. Larynx preservation in pyriform sinus cancer: preliminary results of a European Organization for Research and Treatment of Cancer phase III trial. *J Natl Cancer Inst* 1996; 88: 890–899.
27. Forastiere AA, Goepfert H, Maor M, et al. Concurrent chemotherapy and radiotherapy for organ preservation in advanced laryngeal cancer. *N Engl J Med* 2003; 349: 2091–2098.
28. Colevas AD. Chemotherapy options for patients with metastatic or recurrent squamous cell carcinoma of the head and neck. *J Clin Oncol* 2006; 24: 2644–2652.
29. Forastiere AA, Metch B, Schuller DE, et al. Randomized comparison of cisplatin plus fluorouracil and carboplatin plus fluorouracil versus methotrexate in advanced squamous-cell carcinoma of the head and neck: a Southwest Oncology Group study. *J Clin Oncol* 1992; 10: 1245–1251.
30. Janot F, de Raucourt D, Benhamou E, et al. Randomized trial of postoperative reirradiation combined with chemotherapy after salvage surgery compared with salvage surgery alone in head and neck carcinoma. *J Clin Oncol* 2008; 26: 5518–5523.
31. Salama JK, Vokes EE, Chmura SJ, et al. Long-term outcome of concurrent chemotherapy and reirradiation for recurrent and second primary head-and-neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 2006; 64: 382–391.
32. Rubin Grandis J, Melhem MF, Gooding WE, et al. Levels of TGF-alpha and EGFR protein in head and neck squamous cell carcinoma and patient survival. *J Natl Cancer Inst* 1998; 90: 824–832.
33. Grandis JR, Tweardy DJ. Elevated levels of transforming growth factor alpha and epidermal growth

factor receptor messenger RNA are early markers of carcinogenesis in head and neck cancer. *Cancer Res* 1993; 53: 3579–3584.

34. Karamouzis MV, Grandis JR, Argiris A. Therapies directed against epidermal growth factor receptor in aerodigestive carcinomas. *JAMA* 2007; 298: 70–82.
35. Chung CH, Ely K, McGavran L, et al. Increased epidermal growth factor receptor gene copy number is associated with poor prognosis in head and neck squamous cell carcinomas. *J Clin Oncol* 2006; 24: 4170–4176.
36. Rubin Grandis J, Zeng Q, Tweardy DJ. Retinoic acid normalizes the increased gene transcription rate of TGF-alpha and EGFR in head and neck cancer cell lines. *Nat Med* 1996; 2: 237–240.
37. Vermorken JB, Trigo J, Hitt R, et al. Open-label, uncontrolled, multicenter phase II study to evaluate the efficacy and toxicity of cetuximab as a single agent in patients with recurrent and/or metastatic squamous cell carcinoma of the head and neck who failed to respond to platinum-based therapy. *J Clin Oncol* 2007; 25: 2171–2177.
38. Bonner JA, Harari PM, Giralt J, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *N Engl J Med* 2006; 354: 567–578.
39. Cohen EE, Kane MA, List MA, et al. Phase II trial of gefitinib 250 mg daily in patients with recurrent and/or metastatic squamous cell carcinoma of the head and neck. *Clin Cancer Res* 2005; 11: 8418–8424.
40. Soulieres D, Senzer NN, Vokes EE, et al. Multicenter phase II study of erlotinib, an oral epidermal growth factor receptor tyrosine kinase inhibitor, in patients with recurrent or metastatic squamous cell cancer of the head and neck. *J Clin Oncol* 2004; 22: 77–85.
41. Le Tourneau C, Siu LL. Molecular-targeted therapies in the treatment of squamous cell carcinomas of the head and neck. *Curr Opin Oncol* 2008; 20: 256–263.
42. Mukohara T, Engelman JA, Hanna NH, et al. Differential effects of gefitinib and cetuximab on non-small-cell lung cancers bearing epidermal growth factor receptor mutations. *J Natl Cancer Inst* 2005; 97: 1185–1194.
43. Viloria-Petit A, Crombet T, Jothy S, et al. Acquired resistance to the antitumor effect of epidermal growth factor receptor-blocking antibodies in vivo: a role for altered tumor angiogenesis. *Cancer Res* 2001; 61: 5090–5101.
44. Leer JW, Overgaard J, Heeren G. The European core curriculum on radiotherapy. *Radiother Oncol* 1991; 22: 153–155.
45. Benefield J, Petruzzelli GJ, Fowler S, et al. Regulation of the steps of angiogenesis by human head and neck squamous cell carcinomas. *Invasion Metastasis* 1996; 16: 291–301.
46. Saba NF, Shin DM, Khuri FR. Targeting angiogenesis in head and neck cancer. *Curr Cancer Drug Targets* 2007; 7: 643–649.
47. Choong NW, Mauer AM, Haraf DJ, et al. Phase I trial of erlotinib-based multimodality therapy for inoperable stage III non-small cell lung cancer. *J Thorac Oncol* 2008; 3: 1003–1011.
48. Jones KL, Buzdar AU. Evolving novel anti-HER2 strategies. *Lancet Oncol* 2009; 10: 1179–1187.
49. Wheeler DL, Huang S, Kruser TJ, et al. Mechanisms of acquired resistance to cetuximab: role of HER (ErbB) family members. *Oncogene* 2008; 27: 3944–3956.
50. Hay N, Sonenberg N. Upstream and downstream of mTOR. *Genes Dev* 2004; 18: 1926–1945.
51. Allen GW, Saba C, Armstrong EA, et al. Insulin-like growth factor-I receptor signaling blockade combined with radiation. *Cancer Res* 2007; 67: 1155–1162.
52. Reiter R, Gais P, Jutting U, et al. Aurora kinase A messenger RNA overexpression is correlated with tumor progression and shortened survival in head and neck squamous cell carcinoma. *Clin Cancer Res* 2006; 12: 5136–5141.
53. Elser C, Siu LL, Winquist E, et al. Phase II trial of sorafenib in patients with recurrent or metastatic squamous cell carcinoma of the head and neck or nasopharyngeal carcinoma. *J Clin Oncol* 2007; 25: 3766–3773.
54. Williamson SK, Moon J, Huang CH, et al. Phase II evaluation of sorafenib in advanced and metastatic squamous cell carcinoma of the head and neck: Southwest Oncology Group Study S0420. *J Clin Oncol* 2010; 28: 3330–3335.
55. Haddad RI, Shin DM. Recent advances in head and neck cancer. *N Engl J Med* 2008; 359: 1143–1154.
56. Kao J, Packer S, Vu HL, et al. Phase 1 study of concurrent sunitinib and image-guided radiotherapy followed by maintenance sunitinib for patients with oligometastases: acute toxicity and preliminary response. *Cancer* 2009; 115: 3571–3580.
57. Milano MT, Katz AW, Muhs AG, et al. A prospective pilot study of curative-intent stereotactic body radiation therapy in patients with 5 or fewer oligometastatic lesions. *Cancer* 2008; 112: 650–658.
58. Salama JK, Chmura SJ, Mehta N, et al. An initial report of a radiation dose-escalation trial in patients with one to five sites of metastatic disease. *Clin Cancer Res* 2008; 14: 5255–5259.
59. Holland JF. Breaking the cure barrier 25 years later. *J Clin Oncol* 2008;26(:):; 1575.
60. Choi S, Myers JN. Molecular pathogenesis of oral squamous cell carcinoma: implications for therapy [published correction appears in *J Dent Res* 2008; 87: 191]. *J Dent Res* 2008; 87: 14–32.
61. Lin DT, Subbaramaiah K, Shah JP, et al. Cyclooxygenase-2: a novel molecular target for the prevention and treatment of head and neck cancer. *Head Neck* 2002; 24: 792–799.
62. Kao J, Packer SH, Teng M, et al. Phase I trial of concurrent erlotinib, celecoxib, and reirradiation for recurrent head and neck cancer. *J Clin Oncol* 2010; 28((suppl 15S)):abstract 5561.
63. Kao J, Sikora AT, Fu S. Dual EGFR and COX-2 inhibition as a novel approach to targeting head and neck squamous cell carcinoma. *Curr Cancer Drug Targets* 2009; 9: 931–937.
64. Gupta GP, Nguyen DX, Chiang AC, et al. Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. *Nature* 2007; 446: 765–770.
65. Fukumura D, Kashiwagi S, Jain RK. The role of nitric oxide in tumour progression. *Nat Rev Cancer* 2006; 6: 521–534.
66. Ekmekcioglu S, Tang CH, Grimm EA. NO news is not necessarily good news in cancer. *Curr Cancer Drug Targets* 2005; 5: 103–115.
67. Gallo O, Fabbroni V, Sardi I, et al. Correlation between nitric oxide and cyclooxygenase-2 pathways in head and neck squamous cell carcinomas. *Biochem Biophys Res Commun* 2002; 299: 517–524.

68. Brennan PA, Dennis S, Poller D, et al. Inducible nitric oxide synthase: correlation with extracapsular spread and enhancement of tumor cell invasion in head and neck squamous cell carcinoma. *Head Neck* 2008; 30: 208–214.
69. Tinker AC, Wallace AV. Selective inhibitors of inducible nitric oxide synthase: potential agents for the treatment of inflammatory diseases? *Curr Top Med Chem* 2006; 6: 77–92.
70. Sikora AG, Gelbard A, Davies MA, et al. Targeted inhibition of inducible nitric oxide synthase inhibits growth of human melanoma in vivo and synergizes with chemotherapy. *Clin Cancer Res* 2010; 16: 1834–1844.
71. Shang ZJ, Li ZB, Li JR. In vitro effects of nitric oxide synthase inhibitor L-NAME on oral squamous cell carcinoma: a preliminary study. *Int J Oral Maxillofac Surg* 2006; 35: 539–543.
72. Grandis JR, Pietenpol JA, Greenberger JS, et al. Head and neck cancer: meeting summary and research opportunities. *Cancer Res* 2004; 64: 8126–8129.
73. Aguirre-Ghiso JA. Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer* 2007; 7: 834–846.
74. Sang L, Coller HA, Roberts JM. Control of the reversibility of cellular quiescence by the transcriptional repressor HES1. *Science* 2008; 321: 1095–1100.
75. Schmidt-Kittler O, Ragg T, Daskalakis A, et al. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc Natl Acad Sci U S A* 2003; 100: 7737–7742.
76. Aguirre-Ghiso JA, Estrada Y, Liu D, et al. ERK(MAPK) activity as a determinant of tumor growth and dormancy; regulation by p38(SAPK). *Cancer Res* 2003; 63: 1684–1695.
77. Pan Y, Haines DS. Identification of a tumor-derived p53 mutant with novel transactivating selectivity. *Oncogene* 2000; 19: 3095–3100.
78. Erber R, Conradt C, Homann N, et al. TP53 DNA contact mutations are selectively associated with allelic loss and have a strong clinical impact in head and neck cancer. *Oncogene* 1998; 16: 1671–1679.
79. Kuss I, Hathaway B, Ferris RL, et al. Decreased absolute counts of T lymphocyte subsets and their relation to disease in squamous cell carcinoma of the head and neck. *Clin Cancer Res* 2004; 10: 3755–3762.
80. Lopez-Albaitero A, Nayak JV, Ogino T, et al. Role of antigen-processing machinery in the in vitro resistance of squamous cell carcinoma of the head and neck cells to recognition by CTL. *J Immunol* 2006; 176: 3402–3409.
81. Ferris RL, Hunt JL, Ferrone S. Human leukocyte antigen (HLA) class I defects in head and neck cancer: molecular mechanisms and clinical significance. *Immunol Res* 2005; 33: 113–133.
82. Miyazawa M, Ohsawa R, Tsunoda T, et al. Phase I clinical trial using peptide vaccine for human vascular endothelial growth factor receptor 2 in combination with gemcitabine for patients with advanced pancreatic cancer. *Cancer Sci* 2010; 101: 433–439.
83. Hoffmann TK, Bier H, Whiteside TL. Targeting the immune system: novel therapeutic approaches in squamous cell carcinoma of the head and neck. *Cancer Immunol Immunother* 2004; 53: 1055–1067.
84. Leibowitz MS, Nayak JV, Ferris RL. Head and neck cancer immunotherapy: clinical evaluation. *Curr Oncol Rep* 2008; 10: 162–169.
85. To WC, Wood BG, Krauss JC, et al. Systemic adoptive T-cell immunotherapy in recurrent and metastatic carcinoma of the head and neck: a phase 1 study. *Arch Otolaryngol Head Neck Surg* 2000; 126: 1225–1231.
86. Vujanovic NL, Yasumura S, Hirabayashi H, et al. Antitumor activities of subsets of human IL-2-activated natural killer cells in solid tissues. *J Immunol* 1995; 154: 281–289.
87. Frumento G, Piazza T, Di Carlo E, et al. Targeting tumor-related immunosuppression for cancer immunotherapy. *Endocr Metab Immune Disord Drug Targets* 2006; 6: 233–237.
88. Kim R, Emi M, Tanabe K, et al. Tumor-driven evolution of immunosuppressive networks during malignant progression. *Cancer Res* 2006; 66: 5527–5536.
89. Marigo I, Dolcetti L, Serafini P, et al. Tumor-induced tolerance and immune suppression by myeloid derived suppressor cells. *Immunol Rev* 2008; 222: 162–179.
90. Mougiakakos D, Choudhury A, Lladser A, et al. Regulatory T cells in cancer. *Adv Cancer Res* 2010; 107: 57–117.
91. Dolcetti L, Marigo I, Mantelli B, et al. Myeloid-derived suppressor cell role in tumor-related inflammation. *Cancer Lett* 2008; 267: 216–225.
92. Gabrilovich DI, Nagaraj S. Myeloid-derived suppressor cells as regulators of the immune system. *Nat Rev Immunol* 2009; 9: 162–174.
93. Ozao-Choy J, Ma G, Kao J, et al. The novel role of tyrosine kinase inhibitor in the reversal of immune suppression and modulation of tumor microenvironment for immune-based cancer therapies. *Cancer Res* 2009; 69: 2514–2522.
94. Kao J, Ko EC, Eisenstein S, et al. Targeting immune suppressing myeloid-derived suppressor cells in oncology. *Crit Rev Oncol Hematol* 2010; doi:10.1016/j.critrevonc.2010.02.004.
95. Wirth LJ, Haddad RI, Lindeman NI, et al. Phase I study of gefitinib plus celecoxib in recurrent or metastatic squamous cell carcinoma of the head and neck. *J Clin Oncol* 2005; 23: 6976–6981.
96. Movva S, Verschraegen C. The monoclonal antibody to cytotoxic T lymphocyte antigen 4, ipilimumab (MDX-010), a novel treatment strategy in cancer management. *Expert Opin Biol Ther* 2009; 9: 231–241.
97. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma [published correction appears in *N Engl J Med* 2010;363:1290]. *N Engl J Med* 2010; 363: 711–723.
98. Lee SC, Lopez-Albaitero A, Ferris RL. Immunotherapy of head and neck cancer using tumor antigen-specific monoclonal antibodies. *Curr Oncol Rep* 2009; 11: 156–162.
99. Roda JM, Joshi T, Butchar JP, et al. The activation of natural killer cell effector functions by cetuximab-coated, epidermal growth factor receptor positive tumor cells is enhanced by cytokines. *Clin Cancer Res* 2007; 13: 6419–6428.
100. Bibeau F, Lopez-Crapez E, Di Fiore F, et al. Impact of Fcγ RIIa-Fcγ RIIIa polymorphisms and KRAS mutations on the clinical outcome of patients with

metastatic colorectal cancer treated with cetuximab plus irinotecan. *J Clin Oncol* 2009; 27: 1122–1129.

101. Adili A. Robot-assisted orthopedic surgery. *Semin Laparosc Surg* 2004; 11: 89–98.
102. Hollands CM, Dixey LN. Applications of robotic surgery in pediatric patients. *Surg Laparosc Endosc Percutan Tech* 2002; 12: 71–76.
103. Boehm DH, Arnold MB, Detter C, et al. Incorporating robotics into an open-heart program. *Surg Clin North Am* 2003; 83: 1369–1380.
104. Weinstein GS, O'Malley BW Jr, Desai SC, et al. Transoral robotic surgery: does the ends justify the means? *Curr Opin Otolaryngol Head Neck Surg* 2009; 17: 126–131.
105. Hockstein NG, O'Malley BW Jr, Weinstein GS. Assessment of intraoperative safety in transoral robotic surgery. *Laryngoscope* 2006; 116: 165–168.
106. Machtay M, Perch S, Markiewicz D, et al. Combined surgery and postoperative radiotherapy for carcinoma of the base of radiotherapy for carcinoma of the base of tongue: analysis of treatment outcome and prognostic value of margin status. *Head Neck* 1997; 19: 494–499.
107. O'Malley BW Jr, Weinstein GS, Hockstein NG. Transoral robotic surgery (TORS): glottic microsurgery in a canine model. *J Voice* 2006; 20: 263–268.
108. Hockstein NG, Weinstein GS, O'Malley Jr BW. Maintenance of hemostasis in transoral robotic surgery. *ORL J Otorhinolaryngol Relat Spec* 2005; 67: 220–224.
109. Weinstein GS, O'Malley BW Jr, Hockstein NG. Transoral robotic surgery: supraglottic laryngectomy in a canine model. *Laryngoscope* 2005; 115: 1315–1319.
110. Weinstein GS, O'Malley BW Jr, Snyder W, et al. Transoral robotic surgery: radical tonsillectomy. *Arch Otolaryngol Head Neck Surg* 2007; 133: 1220–1226.
111. Weinstein GS, O'Malley BW Jr, Snyder W, et al. Transoral robotic surgery: supraglottic partial laryngectomy. *Ann Otol Rhinol Laryngol* 2007; 116: 19–23.
112. O'Malley BW Jr, Weinstein GS, Snyder W, et al. Transoral robotic surgery (TORS) for base of tongue neoplasms. *Laryngoscope* 2006; 116: 1465–1472.
113. Mukhija VK, Sung CK, Desai SC, et al. Transoral robotic assisted free flap reconstruction. *Otolaryngol Head Neck Surg* 2009; 140: 124–125.
114. Desai SC, Sung CK, Jang DW, et al. Transoral robotic surgery using a carbon dioxide flexible laser for tumors of the upper aerodigestive tract. *Laryngoscope* 2008; 118: 2187–2189.
115. Genden EM, Desai S, Sung CK. Transoral robotic surgery for the management of head and neck cancer: a preliminary experience. *Head Neck* 2009; 31: 283–289.
116. Desai SC, Sung CK, Genden EM. Transoral robotic surgery using an image guidance system. *Laryngoscope* 2008; 118: 2003–2005.

SPECIAL FEATURE

Discovering Neuroprotection in Parkinson's Disease, or Getting to Haphazard

Karl Kieburtz, MD, MPH

University of Rochester Medical Center, Rochester, NY

OUTLINE

Serendipitous Versus Rational Pathways in Drug Development
Rational Drug Development in Parkinson's Disease
After All, the Value of Serendipity

ABSTRACT

Despite significant research efforts, a therapy to slow or halt the progression of Parkinson's disease ("neuroprotection") remains elusive. The discovery of such a therapy will likely require important observations from a number of perspectives: basic science investigation, clinical research, and careful observation in medical practice. Any possible insights will require confirmation in rigorous clinical trial testing. Combining the useful insights from all perspectives may be the most promising approach to discovering a neuroprotective therapy. ***Mt Sinai J Med 77:700–706, 2010.*** © *2010 Mount Sinai School of Medicine*

Key Words: Clinical trials, neuroprotection, Parkinson's disease.

Parkinson's disease (PD) is a progressive neurodegenerative disorder. Clinical deficits have been described in a number of areas, but the classical description focuses on motor disturbances including resting tremor, slowness of movement (bradykinesia), and increase in muscle tone (rigidity). In addition to these classic motor problems, complex deficits appear over time including freezing, gait dysfunction, postural instability, and falling. Disorders of mood, cognition, autonomic function, and sleep have also been described in PD. All of these features are progressive and eventually lead to significant disability.

Medications are available for the treatment of PD, but these largely improve the classic motor features and may have no effect on, or may even worsen, other aspects of PD. Based on the limited therapeutic options available and the progressive nature of the disease, clinicians and researchers are keenly interested in a treatment that would either slow or stop the progression of this disorder. These treatments have been called "neuroprotective" or "disease-modifying." Although neuroprotection is something that can likely only be truly demonstrated in laboratory animals, this is the most commonly used term for such therapies and will be used for the rest of this article.

> *Based on the limited therapeutic options available and the progressive nature of Parkinson's disease (PD), clinicians and researchers are keenly interested in a treatment that would either slow or stop the progression of this disorder.*

Considerable research efforts over the last 2 decades, at both the preclinical and clinical stage, have focused on finding a neuroprotective treatment for PD. To date, these efforts have not translated

Address Correspondence to:

Karl Kieburtz
University of Rochester
Medical Center
Rochester, NY
Email: karl.kieburtz@ctcc.rochester.edu

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20219

© 2010 Mount Sinai School of Medicine

into a therapy that unequivocally slows or stops the progression of PD. Given the decades of time, millions of dollars, and countless hours of patient, family, and researcher time, it may seem surprising that this effort has not yet yielded a success. On the other hand, discovery of novel therapeutics is often not such a straightforward path.

SERENDIPITOUS VERSUS RATIONAL PATHWAYS IN DRUG DEVELOPMENT

The development of novel and sometimes very potent therapeutics has often followed a serendipitous path. Take, for example, the phosphodiesterase inhibitor sildenafil citrate. This medication was originally developed as an antianginal agent. Unfortunately, it was not found to be effective in trials for the treatment of angina. One reasonable post-hoc question asked whether these failures were due to individuals in the trial not actually taking their medication. In attempting to discover the research subjects' compliance with the experimental medication, the developers found they had a problem actually assessing this issue because many of the male participants in the trial did not return their unused drug supply. Only in pursuit of the reason for this failure to return unused drug supply were the beneficial effects on erectile dysfunction inadvertently discovered. Although this drug failed as an antianginal agent, the serendipitous observation in trials for angina led to subsequent work identifying potential benefits in erectile dysfunction. This story serves to illustrate the observation made by Louis Pasteur in a lecture delivered at the University of Lille in December 1854, where he said "Dans les champs de l'observation le hasard ne favorise que les esprits préparés." This is usually translated to "In the fields of observation chance favors only the prepared mind." There are several aspects of the original French that warrant remark. First, the word that is translated as field, "champs," means field as in a farmer's field, not as an area of work. Pasteur is encouraging would-be scientists to be out "in the field" making important observations. Second, the word that is translated as chance, "hasard," is interesting in its own right. It comes from the Middle French meaning "chance," but it also was derived from the Arabic "az-zahr," meaning "the die," as in dice. In modern English this word is complexed into a more commonly used word: haphazard. This term usually implies chaotic or undirected. The meaning of "hap" in Old Norse is "good luck" or fortune, so a literal translation of "haphazard" might be "good luck chance." In many ways, the discovery of the potential beneficial effects of Viagra (and many other interventions in modern medicine) could be described as "haphazard."

Though serendipitous discovery of potential unexpected therapeutic effects of new compounds is an appealing and even entertaining aspect of drug development, it hardly seems that we can rely on good luck and chance to develop new interventions. The emphasis on basic science discovery in the 20th century was part of an effort to stimulate rational drug development. The basic underlying idea behind such drug development is that with current scientific advances, drug development can be faster and more efficient with pathophysiology-based target selection and drug optimization using medicinal chemistry approaches to structure-function relationships to identify drugs that specifically modulate such important targets in the pathophysiology of major diseases. Although this is an intuitive and appealing hypothesis about how basic science discoveries can translate into more effective medication, the actual efficacy of this approach has been somewhat underwhelming.

For example, between 1995 and 2005, funding for US neuroscience research from both the pharmaceutical industry and federal government sources increased from approximately \$5 billion a year to > \$14 billion a year (Figure 1).[1] This significant increase in investment, one would imagine, should lead to a surge in identification of therapies. However, over that same time, the number of new molecular entities approved as treatments by the Food and Drug Administration stayed constant, or actually began to fall. It might be argued that the time course to see results of such an investment in research takes longer than a decade; still, one might have anticipated at least the beginning of a response in terms of new medications identified.

In response to the apparent difficulties in this rational developmental pathway, the National Institutes of Health developed a policy of trying to re-engineer the clinical research process. One key element of this was the transformation of the General Clinical Research Centers into a national system of Clinical Translational Science Centers, which would provide programmatic support for bench-to-bedside translation of basic science discoveries. This remains a compelling hypothesis, but we have yet to see whether specific programmatic support for translational development, coupled with substantial basic science investment, can actually yield increases in new therapies, particularly for neurodegenerative diseases like PD.

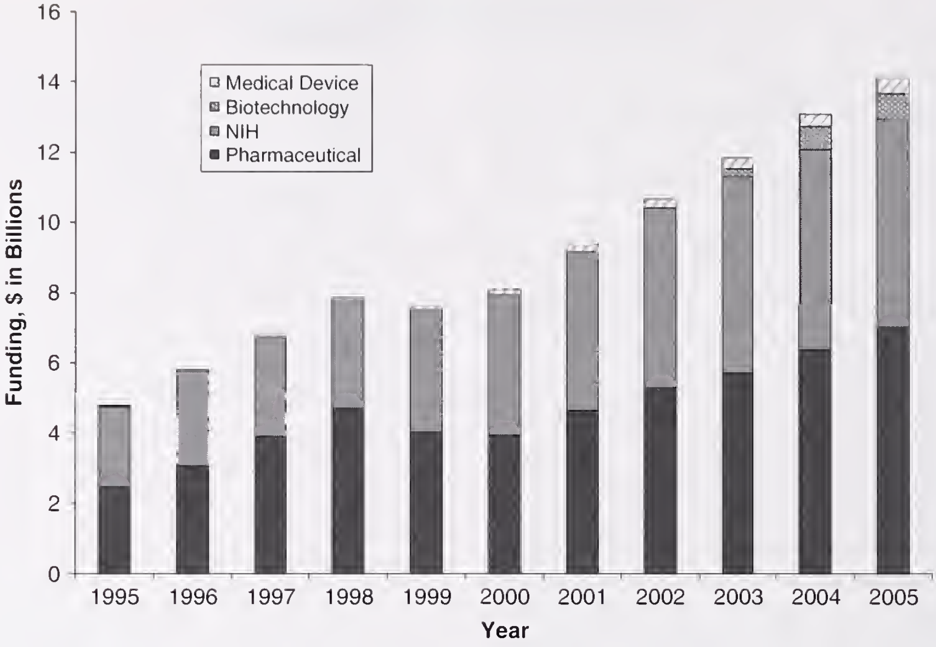


Fig 1. Funding for neuroscience research and development by source, 1995–2005. Striped bars, medical devices; dotted bars, biotechnology; gray bars, NIH; black bars, pharmaceutical industry. **Abbreviations:** NIH, National Institutes of Health. This material is reproduced with permission of John Wiley and Sons, Inc.

RATIONAL DRUG DEVELOPMENT IN PARKINSON'S DISEASE

How well has rational drug development worked for PD? James Parkinson developed the original clinical description of PD almost 200 years ago. It was not until the latter part of the 19th century that the clinical features of PD were found to be associated with substantia nigra and other pigmented brain stem nuclei degeneration. Not until the late 20th century was this substantia nigra degeneration associated with deficiency of striatal dopamine, an observation that led to the Nobel Prize being awarded to several of the individuals involved in these discoveries. The observation of deficient striatal dopamine led to the hypothesis of dopaminergic replacement with the oral precursor levodopa. Although initial trials were plagued with difficulties administering an effective and tolerable dosage, the eventual addition of dopa decarboxylase inhibitors led to effective treatment, one that remains the most potent treatment for PD to this day. Unfortunately, the elapsed time from the observation of striatal degeneration to the initiation of clinical trials with dopaminergic replacement was at least 75 years. This is clearly a time scale that is too slow for our current expectations of therapeutic developments.

There are examples of rational therapeutic development in other areas that have been more expeditious. The observation that chronic myelogenous leukemia is associated with the Philadelphia chromosome, and that this is a translocation that leads to aberrant tyrosine kinase activity, spurred a development program based on looking for tyrosine kinase inhibitors. By searching for molecules with this chemical action, eventually the drug imatinib was identified, following structure-function relationship improvement in possible tyrosine kinase inhibitors. This drug was found to be an effective treatment for chronic myelogenous leukemia and is currently one of the approved treatments. The elapsed time from the observation about tyrosine kinase activity to approval of the drug was on the order of 20 years. This is probably in keeping with expectations of 21st-century drug development. Tyrosine kinase inhibitors have gone on to have some interesting serendipitous extensions of efficacy. For example, another tyrosine kinase inhibitor, gefitinib, was studied for non–small cell lung cancer (NSCLC). Unfortunately, it was not found to be effective in a global efficacy study, although there was apparent efficacy in a Japanese subset of the overall study. The reasons for this were initially unclear. However, subsequent observations were made in individuals with NSCLC that there were specific mutations in the epidermal growth factor receptor that led to responsiveness to this drug.[2] In subsequent trials of individuals with NSCLC, gefitinib has been demonstrated superior to standard therapy only in subjects who carry the mutations, and is an inferior therapy in those who do not carry such epidermal growth factor receptor mutations.[3]

Have we had such brilliant and exciting insights to neuroprotective therapy for PD? Regrettably, not yet. Observations with a series of compounds that seemed to provide neuroprotection in preclinical cellular or animal models of PD, by mechanisms of apoptosis or kinase inhibition, could not be replicated in subsequent clinical trials.[4–6] Indeed, in trials of the 2 agents, there was even a suggestion that they hastened disease progression. It is startling to think that we could actually bring forward interventions to large efficacy trials in PD involving hundreds of patients that turn out to have no effect on disease progression. How might this be so?

Currently we think of the experimental translational therapeutic pathway as a linear process leading from in vitro experiments to in vivo whole animal experiments in relevant models, that then go on to animal safety experiments leading to initial human testing (Table 1).[7] These initial human studies then go on to proof of concept, early studies in humans, and eventually on to confirmatory human efficacy studies that lead to the approval of medications. However, animal models may not be predictive of human diseases. For example, producing rodents or primates that are parkinsonian via intoxication with 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) has provided an excellent model to test treatments to improve PD symptoms. Regrettably, agents that limit MPTP damage in animal models have not been neuroprotective in human testing. Better animal models (perhaps genetic) for testing of potential neuroprotective therapies are urgently needed; however, the experience with stroke models informs us that even excellent animal models are not strongly predictive of the human response.

Currently we think of the experimental translational therapeutic pathway as a linear process leading from in vitro experiments to in vivo whole animal experiments in relevant models, that then go on to animal safety experiments leading to initial human testing.

Animal models may not be predictive of human diseases. For example, producing rodents or primates that are parkinsonian via neurointoxication with 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) has provided an excellent model to test treatments to improve Parkinson's disease symptoms. Regrettably, agents that limit MPTP damage in animal models have not been neuroprotective in human testing.

Table 1. *The Experimental Translational Therapeutic Pathway. This material is reproduced with permission of John Wiley and Sons, Inc.*

In vitro (single cell, cell culture) experiments to define mechanism and candidate interventions
↓
In vivo (whole animal) experiments to assess efficacy in relevant models
↓
Animal experiments for safety
↓
Drug activity studies
↓
IND application to the FDA (in the US)
↓
Initial human testing (Phase I)
↓
Exploratory (proof of concept) human efficacy and safety experiments (Phase II)
↓
Confirmatory human efficacy and safety experiments (Phase III)
↓
Effectiveness in medical practice (Postmarketing surveillance)

More fundamentally, we often lose sight of testing whether interventions actually modulate the biological target of interest in humans. Dosages for further clinical study are often chosen based on the tolerability or on their apparent efficacy in small error-prone, early-phase trials. Dosages and durations are rarely chosen based on quantitative assessment of a drug's ability to enter the physiologic compartment of interest in sufficient concentration to modulate the biologic target in question. Beyond issues of dosage, the entry criteria for a PD study may be intentionally or inadvertently selecting populations that are slower progressors or more resistant to treatment effects. Alternatively, clinical trial designs may be inappropriate for the early-stage development or may use insensitive or incorrect endpoints of drug effects. Lastly, we are currently lacking good measures of "disease activity" (eg, biomarkers) that could aid in evaluating new therapies in both animal and human studies. Large biomarker projects planned by the Michael J. Fox Foundation and the National Institute of Neurological Disorders and Stroke may help address this need. All these complexities of clinical-trial design factor into our current failure to identify a clear-cut neuroprotective therapy in PD.

We are also currently lacking good measures of "disease activity" (eg, biomarkers) that could aid in evaluating new therapies in both animal and human studies. Large biomarker projects planned by the Michael J. Fox Foundation and the National Institute of Neurological Disorders and Stroke may help address this need.

How, then, are we to advance to the discovery of novel treatments for slowing PD progression? We can neither abandon basic science nor the value of clinical observations, even ones in a single patient. How can basic science inform us? A recent example was the finding that the calcium channel blocker isradipine could modulate the firing patterns of dopaminergic neurons.[8] This discovery led to the experiments that assessed the drug's ability to protect mouse neurons from toxins such as rotenone and MPTP. These experiments showed that isradipine was effective in that regard and also had behavioral improvement in the animal model. Based on these initial basic science investigations, as well as supportive epidemiologic data, a clinical trial of isradipine has been started in individuals with early PD (ClinicalTrials.gov Identifier: NCT00909545). Both the safety and tolerability will be assessed, as well as an examination of the impact on the progression of PD signs and symptoms over 1 year. A rigorous trial design will allow for an assessment of whether it is worthwhile to continue investing in the study of isradipine or to drop it in favor of more favorable compounds.

Another source of data is from epidemiologic investigations. Population-based assessments of uric acid levels and the risk of PD have found that individuals with higher plasma urate have a lower risk of PD.[8,9] More recently, observations were made using the data from one of the failed clinical trials alluded to above; they found that high serum urate was also associated with a slower rate of progression than those with manifest PD.[10,11] This is the first time that a serum biomarker was found to be associated with a rate of PD progression. This initial observation was confirmed in a second population of PD patients, the Deprenyl and Tocopherol Antioxidative Therapy of Parkinsonism (DATATOP) cohort, using samples stored from that much earlier clinical trial. Elevation of serum urate is not something to be approached without caution, as a high serum urate is also associated with increased cardiovascular mortality, not only a decreased risk of PD. Based on these observations, a clinical trial of inosine is in progress to see whether serum urate can be safely elevated in individuals with PD, and again whether there is any evidence of a favorable impact on the progression of disease over 1 year (ClinicalTrials.gov Identifier: NCT00833690).

AFTER ALL, THE VALUE OF SERENDIPITY

Lastly, the value of serendipity cannot be overlooked. Serendipitous observations can be made both in basic science experiments as well as in clinical studies. The

The value of serendipity cannot be overlooked. Though initial observations from single or multiple case studies cannot be used as high-quality evidence of efficacy, they are useful for hypothesis generation. Hypotheses based on serendipitous observations can then be tested by *rigorous clinical research techniques such as randomized clinical trials.*

observation that lithium may be beneficial for the treatment of manic-depressive disorder arose from a single investigator using lithium to increase the solubility of urate in an experiment involving guinea pigs. He was trying to test the hypothesis that manic-depressive fluctuations were related to changes in serum urate, but inadvertently found that adding lithium to the solution had a beneficial effect in his animal model. Serendipitous observations also have been important in PD. The original discovery of the effectiveness of amantadine was derived from Parkinson's individuals being treated for influenza. Similarly, the beneficial effects of beta-blockers for essential tremor were inadvertently discovered as these drugs were introduced for the treatment of hypertension and cardiovascular disease. Initial observations from single or multiple case studies cannot be used as high-quality evidence of efficacy. Quite on the contrary, they are useful for hypothesis generation. Hypotheses based on serendipitous observations can then be tested by rigorous clinical research techniques such as randomized clinical trials. The great value of serendipitous observations is that they are based on human clinical events, and do not even require an understanding of the potential mechanism by which these effects are produced. Thus serendipitous observations provide a complementary technique to basic science, target-based therapeutic development, and to observations made in epidemiologic studies. With time, the observations made in serendipitous studies, if verified by subsequent experimentation, should lead to the better identification of specific targets and the optimization of therapeutic interventions.

In summary, there are many routes to discovery. Some of these are more rational and mechanism-based. Others are based on clinical observation and the association of what may initially appear to be random and unrelated events. Individuals with a mind open to discovery, with little prejudice as to the relative merits of these many routes of discovery, will be best placed to capitalize on these early observations. I think this is the point that Pasteur was making in his seminal observation about "the prepared mind." Clearly, mechanism of action should be understood eventually if we are to optimize interventions and their effectiveness, but researchers need to maintain a certain humility and acknowledge their ignorance so that when observations that seem counter to their strongly held beliefs emerge, they are open to possibilities that challenge the current paradigms. A broad acceptance of such possibilities, coupled with rigorous subsequent testing and replication, is likely the best path for discovering a neuroprotective therapy in PD. This path will likely require a lively collaboration among astute clinicians, well-trained clinical researchers, and cutting-edge bench scientists who work together to find potent new treatments.

A broad acceptance of such possibilities, coupled with rigorous subsequent testing and replication, is likely the best path for discovering a neuroprotective therapy in PD. This path will likely require a lively collaboration among astute clinicians, well-trained clinical researchers, and cutting-edge bench scientists who work together to find potent new treatments.

DISCLOSURES

Potential conflict of interest: Dr. Kieburtz has received research grant support from the National Institutes of Health (NEI, NINDS, NIA, NICHD), the Michael J Fox Foundation, Medivation, Neurosearch, and Pfizer; serves as a consultant to The United States Food and Drug Administration, the Veteran's Administration and the National Institutes of Health (NINDS); Abbott, Biogen Idec, Boehringer Ingelheim, Ceregene, EMD Serono, FoldRx, Genzyme, Impax, Ipsen, Isis, Lilly, Lundbeck, Merz , Novartis, Orion, Otsuka, Phytopharm, Prestwick, Schering-Plough, Sienna Biotech, Synosia, Solvay, Teva, UCB Pharma, Xenoport; and provides legal consulting to Pfizer and the Welding Rod Litigation Defendants.

REFERENCES

1. Dorsey ER, Vitticore P, De Roulet J, et al. Financial anatomy of neuroscience research. *Ann Neurol* 2006; 60: 652–659.
2. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor

underlying responsiveness of non-small-cell lung cancer to gefitinib. *N Engl J Med* 2004; 350: 2129–2139.
3. Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009; 361: 947–957.
4. Olanow CW, Schapira AH, LeWitt PA, et al. TCH346 as a neuroprotective drug in Parkinson's disease: a double-blind, randomised, controlled trial. *Lancet Neurol* 2006; 5: 1013–1020.
5. Parkinson Study Group PRECEPT Investigators. Mixed lineage kinase inhibitor CEP-1347 fails to delay disability in early Parkinson disease. *Neurology* 2007; 69: 1480–1490.
6. Kieburtz K, Olanow CW. Translational experimental therapeutics: the translation of laboratory-based discovery into disease-related therapy. *Mt Sinai J Med* 2007; 74: 7–14.
7. Chan CS, Guzman JN, Ilijic E, et al. 'Rejuvenation' protects neurons in mouse models of Parkinson's disease. *Nature* 2007; 447: 1081–1086.
8. De Lau LM, Koudstaal PJ, Hofman A, et al. Serum uric acid levels and the risk of Parkinson disease. *Ann Neurol* 2005; 58: 797–800.
9. Weisskopf MG, O'Reilly E, Chen H, et al. Plasma urate and risk of Parkinson's disease. *Am J Epidemiol* 2007; 166: 561–567.
10. Schwarzschild MA, Schwid SR, Marek K, et al. Serum urate as a predictor of clinical and radiographic progression in Parkinson disease. *Arch Neurol* 2008; 65: 716–723.
11. Ascherio A, LeWitt PA, Xu K, et al. Urate as a predictor of the rate of clinical decline in Parkinson disease. *Arch Neurol* 2009; 66: 1460–1468.

SPECIAL FEATURE

A Cross-Cultural Perspective on Irritable Bowel Syndrome

Charles D. Gerson, MD[1] and Mary-Joan Gerson, PhD[1,2]

[1]Division of Gastroenterology, Mount Sinai School of Medicine, New York, NY
[2]New York University Post-doctoral Program in Psychotherapy and Psychoanalysis, New York, NY

OUTLINE

A Cross-Cultural Perspective
Our Cross-Cultural Research
Views of Irritable Bowel Syndrome Around the Globe
  Sendai, Japan
  Kolkata, India
  Mexico City, Mexico
  Hong Kong
  Bucharest, Romania
  Bari, Italy
Conclusion

ABSTRACT

Irritable bowel syndrome is a functional gastrointestinal illness, defined by symptoms. Irritable bowel syndrome has been described as a biopsychosocial condition, in which colonic dysfunction is affected by psychological and social factors. As a result of this unusual constellation, irritable bowel syndrome may be subject to cultural variables that differ in different parts of the globe. In this article, we describe some of the ways in which irritable bowel syndrome may be experienced differently, depending on local belief systems, psychological pressures, acceptance or resistance to a mind-body paradigm, and breakdown in support or relationship structure. Examples are given in which irritable bowel syndrome investigators from countries around the world describe various aspects of the syndrome that may affect the illness experience of their patients. We describe our own research studies that have demonstrated possible adverse effects on disease severity from relationship conflict, attribution of symptoms to physical rather than emotional cause, and the belief that irritable bowel syndrome is enduring and mysterious. Also described is our finding that symptom patterns may differ significantly between different geographic locations. Finally, we discuss the importance of "cultural competence" on the part of healthcare professionals in regard to caring for patients of diverse cultural backgrounds. ***Mt Sinai J Med 77:707–712, 2010.*** © *2010 Mount Sinai School of Medicine*

Address Correspondence to:

Charles D. Gerson
Division of Gastroenterology
Mount Sinai School of Medicine
New York, NY
Email: cgerson@yahoo.com

Key Words: cross-cultural, cultural competence, irritable bowel syndrome.

Most illnesses occur on a global scale. Many of them manifest themselves the same way in different geographic locations. Although leading American medical journals now contain a much higher percentage of research articles from other countries than in the past, there generally remains a Western-centric description of disease. This ignores the possibility that there are some illnesses that may manifest themselves differently in different places, depending on a number of local variables such as belief systems, psychosocial factors, diet, and socioeconomic changes. Recent studies suggest that irritable bowel syndrome (IBS), a functional gastrointestinal (GI) disorder defined by symptoms of abdominal pain or discomfort and a disordered bowel habit, may be such an illness.

A CROSS-CULTURAL PERSPECTIVE

> *Irritable bowel syndrome is a prototypical biopsychosocial condition affected by local beliefs and traditional belief symptoms.*

The GI tract functions normally through a myriad of finely tuned interactions between smooth muscle,

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI:10.1002/msj.20218

© 2010 Mount Sinai School of Medicine

the autonomic and enteric nervous system, local neurotransmitters, and hormones. This function can be perturbed, at least in part, by emotional factors, leading to a series of functional syndromes including noncardiac chest pain, nonulcer dyspepsia, and IBS. This had led to common expressions such as "gut feelings," "butterflies in the stomach," and "scared s...less," folk wisdom supported by scientific study.

How can local beliefs and traditional belief systems affect IBS, a prototypical biopsychosocial conditon?[1] Kleinman, a pioneer in the area of medical anthropology, says, "Illness behavior is a normative experience governed by cultural rules."[2] Those rules can include how symptoms are perceived, expressed, and coped with. Kleinman gives an example of a Chinese patient suffering from major depression whose symptoms were expressed as weakness, back pain, weight loss, and heaviness in his feet; they were attributed by the patient to sexual excess, but the patient denied any psychological distress. Kleinman also describes Chinese patients who suffered emotionally from the Cultural Revolution but were only able to express this through somatic complaints.[3]

Irritable bowel syndrome is an unusual syndrome in that it is defined by a set of symptoms rather than pathologic findings or biomarkers. Patients complain about abdominal pain, a disordered bowel habit, bloating, distension, and gas, all of which contain a subjective element. Irritable bowel syndrome is considered a mind-body illness with a circular pattern: physical symptoms may affect emotions, and in turn, emotional distress affects symptoms.

Irritable bowel syndrome (IBS) is considered a mind-body illness with a circular pattern: physical symptoms may affect emotions, and in turn, emotional distress affects symptoms.

Is it possible that beliefs may have an impact on the IBS illness experience? We have published several articles that suggest this to be the case. In a survey of IBS patients in 8 countries, to be described in greater detail below, subjects were asked to what degree they attributed their GI symptoms to physical or emotional causation.[4] Those with greater attribution to emotional factors had lower symptom severity, whereas patients who gave greater weight to physical causation had higher symptom severity. One possible explanation of these findings is that cultural resistance to emotional causation may result in increased somatic symptomatology, similar to the Chinese patient described above.

We also found that several negative beliefs–for example, that IBS symptoms will be enduring and that IBS is a mysterious illness (poorly understood)–correlated with greater symptom severity.[5] A sense of mystery surrounding illness may be more injurious in one culture than another. Belief by IBS patients that the condition will endure has also been reported elsewhere and was correlated with another negative belief, that patients have little control over outcome.[6]

What other effects can local variables have on the IBS experience? Irritable bowel syndrome seems to be affected by geographic dislocation, as illustrated by Bedouins living in Israel who were found to have a significantly higher prevalence of IBS after being moved to urban housing compared with Bedouins still living a nomadic existence.[7] Taking a population out of its natural habitat, especially if done forcibly, clearly can lead to various stresses that may affect health. There is a trend upward in IBS prevalence in the developing world, exemplified by 2 studies in Singapore in which IBS prevalence increased from 2.3% to 11% from 1998 to 2004, using the same symptom criteria.[8] In addition to the obvious stresses of urban life, cultural variables may include the change from extended family to nuclear family, with an associated diminution of social or community support. Low social support has been reported as a feature of IBS.[9]

There is a trend upward in IBS prevalence in the developing world, exemplified by 2 studies in Singapore in which IBS prevalence increased from 2.3% to 11% from 1998 to 2004, using the same symptom criteria.

An important cross-cultural area is cultural competence on the part of the health professional, defined as an ability to interact effectively with people of different cultures.

An important cross-cultural area is cultural competence on the part of the health professional, defined as an ability to interact effectively with people of different cultures.[10] For example, if a physician with strong biomedical beliefs is treating a patient with traditional beliefs, there will be a profound disconnect in the treatment approach. In the book *The Spirit Catches You and You Fall Down: A Hmong Child, Her American Doctors, and the Collision of Two Cultures*,[11] Fadiman describes a Hmong child with epilepsy whose family belief was that medical cure must include a ritual sacrifice of a goat. Unresolved conflict with the child's physicians resulted in rejection of Western medical treatment and the death of the child.

We have been actively engaged in IBS cross-cultural research and have also been editing comments by international colleagues for a Functional Brain-Gut Research Group biannual newsletter. In this review, we will describe our own research findings and give some examples of our colleagues' observations. Their descriptions exemplify the kinds of cultural differences that can affect the way that IBS is experienced in different locations around the globe. They also show the potential difficulties that healthcare professionals in the United States may have if they encounter patients from other cultures.

There are some important caveats when comparing different cultures with different languages. One obvious issue is translation of questionnaires. One of our research collaborators, Ami Sperber of Israel, has written extensively about validation of translated questionnaires.[12] We and our colleagues used the back-translation method, considered to be a valid process. Our questionnaires were translated by a bilingual individual into the native language. Then a second bilingual person translated the local language back into English. Where discrepancies occurred between the original text and the back-translated text, the conflict was resolved to the satisfaction of the 2 consultants. We are confident that the back-translation method accounts for important differences in the cultural meaning of subjective distress, such as pain or bloating.

Another issue is the cross-cultural validity of measures. Fortunately, the Quality of Relationship Inventory (QRI), one of our research questionnaires that is most likely to be affected by cultural beliefs, has been used in different populations. In a study that compared the QRI in American students and first-generation Vietnamese students, internal consistency of responses for both groups were compared to each other and to previously published data, and alpha coefficients were highly similar for both groups.[13] We were encouraged by this cross-cultural validation data, but we recognize that application of the QRI to other cultures might be more variable.

OUR CROSS-CULTURAL RESEARCH

We have completed an 8-country survey of IBS around the globe and are currently in the midst of a more extensive 9-country survey. Both are mainly focused on psychosocial variables but also include symptom severity and symptom patterns. Details have been published elsewhere.[4,14]

In the completed study, we collaborated with IBS investigators in Mexico City; Montreal, Canada; London, England; Bari, Italy; Beersheva, Israel; Kolkata, India; and Beijing, China.[4] The psychosocial variables were family relationships, as well as the degree to which IBS patients attribute their symptoms to emotional or physical causality. The major symptom patterns seen in IBS patients (abdominal pain, constipation, diarrhea, and bloating) were compared from country to country.

Thirty patients were surveyed at each site, most of which were tertiary care centers. Symptoms were assessed with a Bowel Symptom Scale that measured the 4 symptoms utilizing a Likert scale. Family dynamics were measured by the QRI and included subscales for family support, depth, and conflict. A mind-body IBS questionnaire was used for emotional or physical attribution. It contained 10 questions each, describing possible physical or psychological etiology of IBS symptoms.

Composite results for the psychosocial variables were quite significant. Severity of symptoms was highly correlated with conflict in the family, whereas support and depth correlated with low symptom severity. Interestingly, attribution of symptoms to physical causality correlated significantly with higher symptom score, and emotional causality correlated with lower symptom score.[4] Our results are consistent with other reports showing that relationship conflict may adversely affect many illnesses.[15]

A survey of irritable bowel syndrome patients in 8 countries revealed that the severity of symptoms was highly correlated with conflict in the family. In addition, a patient's attribution of symptoms to physical causality correlated significantly with higher symptom score, and emotional causality correlated with lower symptom score.

The other arm of this study was a comparison of symptoms between countries.[14] Ours is the first investigation of geographic symptom differences on a global scale. Most cross-cultural research in IBS is restricted to prevalence rates. Interesting differences between countries were noted. Of the 8 countries studied, China had significantly higher diarrhea scores than 5 other countries. It is possible that the intestinal flora in Chinese subjects is different than in subjects in the other countries studied because of local gut contamination. This is relevant, because IBS is known to follow acute gastroenteritis and may be affected by mucosal inflammation. Mexican subjects had significantly higher constipation scores, confirmed in another publication,[16] and this is likely to be related to a high-starch diet.

We are now engaged in a larger study of psychosocial variables involving 500 patients at 10 study sites around the world: New York City and Los Angeles in the United States, plus Mexico, Nicaragua, Germany, Italy, Iran, India, China, and Japan. We are assessing the ways in which attachment style, pain beliefs, and catastrophizing are correlated with IBS symptom severity.

VIEWS OF IRRITABLE BOWEL SYNDROME AROUND THE GLOBE

The following are condensed versions of commentaries submitted by our international colleagues, who were asked to describe IBS in their respective countries.

Sendai, Japan

Dr. Shin Fukudo reports that epidemiological studies of IBS in Japan have shown a similar prevalence to data from the United States, probably reflecting Japan's relatively advanced economic and urbanized status. Belief systems of Japanese people are in a transition state. Traditionally, there has been a strong belief in the connection between mind and body. Certain expressions represent this belief; for example, to say "moving gut" means getting angry, and "showing inside the body" signifies not revealing one's secrets.

However, modern concepts are displacing this holistic belief with a more dichotomized view of mind and body. Physicians have encouraged this change by conceptualizing IBS as either purely organic or purely psychogenic. This is unfortunate for 2 reasons: the rejection of a biopsychosocial model makes it more difficult to help patients; and if patients are viewed as burdened only by psychological distress, they still resist psychological treatment because of the stigma associated with it. Unfortunately, in many of the countries we have studied, referral to mental health professionals is viewed as signifying severe psychiatric disease, an attribution somewhat shifted in our own culture. Secondly, the focus on organic disease rather than functional illness leads to excessive testing and procedures.

In both the United States and Japan, there is a dynamic tension between Western and holistic beliefs. Among patients, the pendulum seems to be slowly turning back to a more holistic approach, but this has been coupled with significant resistance by the medical profession.

Kolkata, India

Our colleague Abhijit Chowdhury describes a unique set of beliefs in his Indian patients. Although India has a strong tradition of Ayurvedic medicine based on humors like bile and soul, this is being supplanted by other beliefs. However, those beliefs cannot be described as modern in the Western sense, as there is a strong prevailing opinion held by patients and physicians that IBS symptoms are caused by giardiasis, without documentation. There is also a strong belief that passage of flatus is critical for symptom relief, not only for IBS, but also for migraine headaches. Though IBS is often associated with anxiety or depression, there is a strong taboo against being labeled as psychologically unstable.

Another interesting cultural aspect related to the care of IBS patients in India concerns gender. It has long been reported that the usual female predominance in IBS is reversed in India, based on several research studies performed in the 1980s and 1990s that are frequently quoted. Dr. Chowdhury explains that the male predominance in prevalence studies is related mainly to the fact that healthcare is primarily accessed by males. In rural areas, where the majority of Indians live, there is a strong gender bias that keeps females in a relatively subservient role, often prevented from accessing the same medical care available to men. With modernization, it is likely that the unusual gender ratio reported in India will change as women have improved access to healthcare.

Mexico City, Mexico

Dr. Max Schmulson reports that emotions and their expression are an important part of Mexican culture. As a result, it is generally accepted that IBS is related, at least in part, to stressful situations, usually involving the family. As one patient stated, "Stress generates my symptoms. I'm stressed because of my husband, who has a very bad character and has fights constantly, for example in restaurants."

The emphasis on family reveals a phenomenon, documented in many studies of family dynamics and illness, that intimate relationships can have a significant effect on health. We frequently encounter this in our patients and feel that recognition and discussion of family interactions around the IBS patient is an important part of the treatment process.

The term "macho" is obviously of Spanish origin and adds a particular flavor to the IBS experience in Mexico. Women are much more open about discussing their GI symptoms than are men, who find this embarrassing and tend to express their concerns about the IBS experience by joking.

Hong Kong

Our colleague Grace Wang has described a very different concept of illness in her IBS patients in Hong Kong than we see in the United States. Chinese patients consider themselves to be intimately connected to their environment. If they are ill, they are out of balance with their environment; they see that balance, for example, in terms of yin and yang, hot and cold, or dry and damp.

As a result, Chinese patients feel an unusual degree of responsibility for their illness and seek to remedy it by re-establishing balance through eating various hot or cold foods that would not make sense to a Western physician. Patients with IBS often seek the help of traditional healers, who may have more to offer than Western medicine because their treatment will be more congruent with the patient's belief system.

Bucharest, Romania

Dan Dumitrascu, an IBS investigator in Romania, describes a very different situation in that country. Romania has gone through a very difficult political history, enduring a suppressive Communist regime for many years until the breakup of the Communist world. While this caused a great deal of stress that may have been manifested, in part, by GI functional symptoms, the government suppressed medical research in this area, and the political use of psychiatry resulted in an aversion to seeking psychological help.

Some IBS research started taking place in the 1990s. Dr. Dumitrascu describes psychosocial stress in his patients, yet there is still an avoidance of psychological help, though people are starting to see psychologists—not psychiatrists, who are associated with political suppression. Probably because of the politicization of mental health treatment, beliefs about IBS vary but tend toward motility disturbance and intestinal infection.

Bari, Italy

Piero Porcelli states, "It is a strong popular belief in Italy (at least in the south of Italy, where I live and work) that somatic discomfort and illness is caused by interpersonal distress, mainly in family and emotional relationships. Family is highly valued in our culture, not only as a place of safety and acceptance, but also as a place of personal identity, so that problems in family relationships and communication can be experienced (at a conscious as well as unconscious level) as a sort of decrease in personal strength, a part of the self that is damaged by the conflicts with relatives." This is an excellent example of the way that community and cultural rules quoted from Kleinman can affect an illness like IBS.

Dr. Porcelli also describes how southern Italy is experiencing a shift from the old family physician with a holistic and empathic approach to the new, more technological physician focused on test results. In a survey he conducted, 30% of patients complained that their symptoms are due to physicians who do not understand their health problems.

CONCLUSION

Through our work, we have demonstrated that it is possible to conduct international research with colleagues in far-reaching locations, using the Internet in what might be called research collaboration without walls. We believe that it is important to realize that the way we see illness manifested in the United States may not be representative of the rest of the world.[17] Global study gives us a more complete picture of an illness and also can help us to improve our treatment of patients who come from distant cultures.

ACKNOWLEDGMENT

This work was supported by the Ping Y. Tai Foundation.

DISCLOSURES

Potential conflict of interest: Nothing to report.

REFERENCES

1. Drossman DA. Gastrointestinal illness and the biopsychosocial model. *J Clin Gastroenterol* 1996; 22: 252–254.
2. Kleinman A. The Illness Narratives: Suffering, Healing and the Human Condition. New York, NY: Basic Books; 1988; 6.
3. Kleinman A. Depression, somatization and the "new cross-cultural psychiatry." *Soc Sci Med* 1977; 11: 3–10.
4. Gerson MJ, Gerson CD, Awad RA, et al. An international study of irritable bowel syndrome: family relationships and mind-body attributions. *Soc Sci Med* 2006; 62: 2838–2847.
5. Gerson MJ, Gerson CD. Relationship between symptom severity and belief systems in irritable bowel syndrome. *Gastroenterology* 2008; 134:(suppl 1): A–418.
6. Rutter CL, Rutter DR. Illness representation, coping and outcome in irritable bowel syndrome (IBS). *Br J Health Psychol* 2002; 7: 377–391.
7. Sperber AD, Friger M, Shvartzman P, et al. Rates of functional bowel disorders among Israeli Bedouins in rural areas compared to those who moved to permanent towns. *Clin Gastroenterol Hepatol* 2005; 3: 342–348.
8. Gwee KA, Wee S, Wong ML, et al. The prevalence, symptom characteristics, and impact of irritable bowel syndrome in an Asian urban community. *Am J Gastroenterol* 2004; 99: 924–931.
9. Jones MP, Wessinger S, Crowell MD. Coping strategies and interpersonal support in patients with irritable bowel syndrome and inflammatory bowel disease. *Clin Gastroenterol Hepatol* 2006; 4: 474–481.
10. Fox RC. Cultural competence and the culture of medicine. *N Engl J Med* 2005; 353: 1316–1319.
11. Fadiman A. *The Spirit Catches You and You Fall Down: A Hmong Child, Her American Doctors and the Collision of Two Cultures*. New York, NY: Farrar, Straus and Giroux; 1997.
12. Sperber AD. Translation and validation of study instruments for cross-cultural research. *Gastroenterology* 2004; 126: S124–S128.
13. Dinh KT, Sarason BR, Sarason IG. Parent-child relationships in Vietnamese immigrant families. *J Fam Psychol* 1994; 8: 471–488.
14. Gerson CD, Gerson MJ, Awad RA, et al. Irritable bowel syndrome: an international study of symptoms in eight countries. *Eur J Gastroenterol Hepatol* 2008; 20: 659–667.
15. Kiecolt-Glazer JK, Newton TL. Marriage and health: his and hers. *Psychol Bull* 2001; 127: 472–503.
16. Schmulson M, López-Alvarenga JC, Vargas J. Clinical differences in irritable bowel syndrome (IBS) subgroups according to Rome III criteria: a preliminary report of the Mexican IBS working group. *Neurogastroenterol Motil* 2008; 20:(suppl 1): 138.
17. Drossman DA, Weinland SR. Commentary: sociocultural factors in medicine and gastrointestinal research. *Eur J Gastroenterol Hepatol* 2008; 20: 593–595.

UNITED STATES POSTAL SERVICE®

Statement of Ownership, Management, and Circulation (All Periodicals Publications Except Requester Publications)

1. Publication Title Mount Sinai Journal of Medicine	2. Publication Number 2 8 4 - 8 8 0	3. Filing Date 10/1/10
4. Issue Frequency 6x's/yr in April, June, Aug, Sept, Nov, Dec	5. Number of Issues Published Annually 6	6. Annual Subscription Price $105.00

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4®)*

Wiley Subscription Services, Inc., 111 River Street, Hoboken, NJ 07030

Contact Person
E. Schmidlchen

Telephone *(Include area code)*
(201) 748-6346

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

Wiley Subscription Services, Inc., 111 River Street, Hoboken, NJ 07030

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*

Wiley Subscription Services, Inc., 111 River Street, Hoboken, NJ 07030

Editor *(Name and complete mailing address)*

Penny A. Asbell, Dept. of Ophthalmology, One Gustave L. Levy Place, Box 1183, New York, NY 10029

Managing Editor *(Name and complete mailing address)*

None

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	**Complete Mailing Address**
Mount Sinai School of Medicine	1 Gustave Levy Place
	New York, NY 10029

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☑ None

Full Name	**Complete Mailing Address**

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*
The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes.
☐ Has Not Changed During Preceding 12 Months
☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

PS Form **3526**, September 2006

13. Publication Title Mount Sinai Journal of Medicine	14. Issue Date for Circulation Data July/August 2010

15. **Extent and Nature of Circulation**			**Average No. Copies Each Issue During Preceding 12 Months**	**No. Copies of Single Issue Published Nearest to Filing Date**
a. Total Number of Copies *(Net press run)*			4289	4100
b. Paid Circulation *(By Mail and Outside the Mail)*	(1)	Mailed Outside-County Paid Subscriptions Stated on PS Form 3541 *(Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)*	2781	2689
	(2)	Mailed In-County Paid Subscriptions Stated on PS Form 3541 *(Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)*	0	0
	(3)	Paid Distribution Outside the Mails Including Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Paid Distribution Outside USPS®	0	0
	(4)	Paid Distribution by Other Classes of Mail Through the USPS (e.g. First-Class Mail®)	0	0
c. Total Paid Distribution *(Sum of 15b (1), (2),(3), and (4))*			2781	2689
d. Free or Nominal Rate Distribution *(By Mail and Outside the Mail)*	(1)	Free or Nominal Rate Outside-County Copies included on PS Form 3541	399	397
	(2)	Free or Nominal Rate In-County Copies Included on PS Form 3541	0	0
	(3)	Free or Nominal Rate Copies Mailed at Other Classes Through the USPS (e.g. First-Class Mail)	0	0
	(4)	Free or Nominal Rate Distribution Outside the Mail *(Carriers or other means)*	0	0
e. Total Free or Nominal Rate Distribution *(Sum of 15d (1), (2), (3) and (4)*			399	397
f. Total Distribution *(Sum of 15c and 15e)* ▶			3180	3086
g. Copies not Distributed *(See Instructions to Publishers #4 (page #3))* ▶			1109	1014
h. Total *(Sum of 15f and g)* ▶			4289	4100
i. Percent Paid *(15c divided by 15f times 100)* ▶			87.45	87.14

16. Publication of Statement of Ownership
☑ If the publication is a general publication, publication of this statement is required. Will be printed in the Nov/Dec 2010 issue of this publication.
☐ Publication not required

17. Signature and Title of Editor, Publisher, Business Manager, or Owner

Elizabeth Konkle, Associate Financial Manager EKonkle

Date
10/1/10

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

PS Form **3526**, September 2006

The Evidence-Based Resource Trusted by Health Professionals

ESSENTIAL EVIDENCE PLUS

Essential Evidence Plus (EE+) is the only point-of-care resource that provides you with the Bottom Line to help your patients. This powerful new e-reference tool includes medical evidence ratings so you know you are getting the best EBM answers for teaching and patient care.

EE+ gives you everything you need to make critical patient care decisions, and includes must-have content, calculators, daily POEM updates, guidelines, and tools that clinicians can rely on for the best answer. It has the latest medical findings pre-screened for validity and relevance to clinical practice, and can be used to search all of the evidence, all of the time, in an instant. **EE+** is packed with great features you won't find with any other comparable product and is the most complete evidence-based medicine resource available!

Unique features include:

- **NEW RESOURCE!** ***Essential Evidence*** featuring 700+ current clinical topics designed to provide evidence-based answers to important questions concerning symptoms, diseases, and treatment that arise in primary care and hospital practice
- 1,900+ diagnostic test calculators
- 1,700+ history and physical test calculators
- 300+ interactive decision support tools and clinical decision rules
- Daily POEM e-mail alerts to help you stay abreast of the latest and most relevant medical literature and 3,500+ Archived POEMs
- 3,300+ Cochrane Abstracts
- EBM Guidelines brought to you by the Finnish Medical Society with 1,000+ practice guidelines, 3,000+ Evidence-Graded Summaries and 950+ high quality photographs
- Derm Expert viewer with 950+ pictures designed to help clinicians diagnose dermatological conditions
- 1,600+ Guideline Summaries
- 1,500+ Commonly used ICD-9 Codes
- Weekly Podcasts
- CME credit availability for qualified searches

Don't take our word for it.
Check out what others are saying
about Essential Evidence Plus!

> "**Essential Evidence Plus** is an important tool that will provide primary care clinicians with access to the most extensive source of evidence-based medical information currently available for use at the point of care."
>
> — Michael Wilkes, MD,
> Professor of Medicine and
> Vice Dean of Education,
> University of California, Davis

> "This database will revolutionize the way librarians work with clinicians. [It] has set the standard for all clinical database designers to surpass if they want to succeed."
>
> — Rick Wallace, Outreach Librarian,
> Eastern Tennessee State University

> "...places highly valid, highly relevant information 'at the fingertips' of clinicians while they practice."
>
> — Journal of the Medical
> Library Association

> "This is the most robust, able medical resource available."
>
> — David Stockwell, MD,
> Children's National Medical Center

Sign up for a free trial or subscribe today!
Group price quotations are also available upon request.

www.essentialevidenceplus.com

ESSENTIAL EVIDENCE PLUS
The EBM Superstore for the Point-of-Care
Search once and find the Bottom Line!

Essential Evidence Plus (EE+) is the most powerful, comprehensive, evidence-based database trusted by health professionals. It is the only e-reference tool that searches many EBM resources at once and is packed with everything you need to make critical patient-care decisions. EE+ includes must-have content, calculators, daily POEM email alerts, guidelines, decision support tools, and more that you can rely on for the best answer.

EE+ is the first place to go for answers to point-of-care questions in record time and includes unlimited access to these world renowned resources — tools you won't find in any other comparable product.

- ✓ **Essential Evidence** — NEW RESOURCE EXCLUSIVELY FROM EE+! Concise, easily digestible summaries for the front line that bring information into practice in one search. Essential Evidence is written by leading physicians and features more than 700 topics, covering the most common conditions, diseases, and procedures. All relevant medical literature is reviewed, evaluated, and then summarized down to the Bottom Line answers you can get in seconds. It also includes medical evidence ratings so you know you are getting the best EBM answers for teaching and patient care, **PLUS** they link to other evidence-based medicine resources within Essential Evidence **Plus**.
- ✓ **Daily POEMs (Patient-Oriented Evidence that Matters)** e-mail alerts and **Archived POEMs** to help you stay abreast of the latest and most relevant medical literature
- ✓ **EBM Guidelines** brought to you by the Finnish Medical Society with practice guidelines, Evidence-Graded Summaries, and high quality photographs
- ✓ **Cochrane Abstracts**
- ✓ **Decision Support Tools and Diagnostic, History and Physical Test Calculators** to guide the most accurate diagnosis, assess risk and probability, select the safest treatments, indicate which tests to order, and help you interpret the results
- ✓ **Practice Guidelines**
- ✓ **ICD-9 Codes**
- ✓ **Number Needed to Treat** (NNT)
- ✓ **Podcasts**
- ✓ **CME** credit available for qualified searches

Check off everything at once! **ESSENTIAL EVIDENCE PLUS** — Your one stop for all the EBM resources you need for better patient care.

Sign up for a free trial or to subscribe today!

INSIDER DISCOUNT

SAVE $10!

Use discount code **SHOPCZ9R** during step 3 of the checkout process!

Group price quotations are available upon request.

www.essentialevidenceplus.com

LEVY LIBRARY, MT. SINAI SCHOOL OF MEDICINE
3 4805 0909018 0

FREE access to cutting-edge content

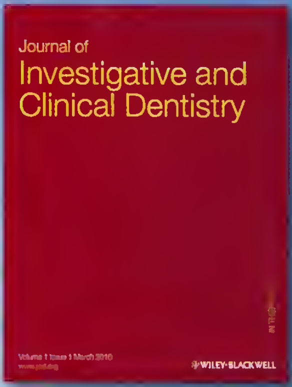
New Journals for 2010

WILEY-BLACKWELL

Clinical and Experimental Neuroimmunology

Global Policy

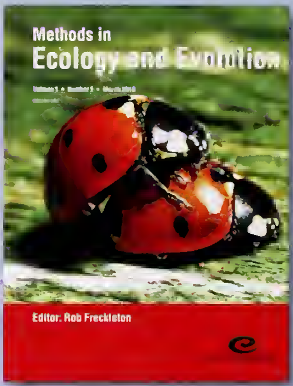
Journal of Diabetes Investigation



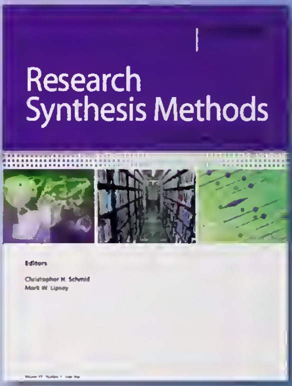
Journal of Investigative and Clinical Dentistry

Journal of Pharmaceutical Health Services Research

Journal of Psychological Issues in Organizational Culture



Methods in Ecology and Evolution



Research Synthesis Methods

WILEY
Interdisciplinary Reviews

WIREs are unique hybrid publications that combine the most powerful features of online reference works with review journals, emphasizing the importance of interdisciplinary collaboration in research and education.

There are three titles launching in 2010:

- Climate Change
- Cognitive Science
- RNA

Register for free online access at:
www.interscience.wiley.com/newjournals

15529

MOUNT SINAI
GUSTAVE L.
AND
JANET W.
LEVY
LIBRARY

MOUNT SINAI
GUSTAVE L.
AND
JANET W.
LEVY
LIBRARY

LEVY LIBRARY, MT. SINAI SCHOOL OF MEDICINE
3 4805 0913381 6